

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:14 : Search time 35 Seconds  
(without alignments)  
41.879 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSGKLSLX 11

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.GeneSeq\_101002:\*

1: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SID52/gcgcdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	95.2	9	21	AA14276
2	40	95.2	9	21	AA191252
3	40	95.2	11	21	AA14273
4	40	95.2	11	23	AAU80306
5	40	95.2	20	21	AA14286
6	40	95.2	29	21	AA191258
7	40	95.2	29	21	AA191259
8	33	78.6	38	22	AB131326
9	33	78.6	38	22	AB136532
10	33	78.6	38	22	AB121873

11	33	78.6	38	22	AA157294	Human brain expres
12	33	78.6	38	22	AA166699	Human bone marrow
13	33	78.6	38	22	AA17511	Peptide #3945 enco
14	33	78.6	38	22	AA130033	Peptide #4070 enco
15	33	78.6	38	22	AA105183	Peptide #3865 enco
16	33	78.6	38	23	AB139314	Human peptide enco
17	33	78.6	192	22	AB161074	Drosophila melanog
18	33	78.6	559	23	AB193132	Herbivoidally activ
19	33	78.6	571	21	AA136317	Arabidopsis thalia
20	32	76.2	76	22	AA107277	Human polypeptide
21	32	76.2	301	22	AA193760	Human protein sequ
22	32	76.2	352	22	AB103657	Novel human diago
23	31	73.8	93	21	AA154297	Human pancreatic c
24	31	73.8	140	21	AA118668	Zea mays protein f
25	31	73.8	165	21	AA142267	Human ORF2031
26	31	73.8	299	21	AA134445	Arabidopsis thalia
27	31	73.8	312	21	AA134444	Arabidopsis thalia
28	31	73.8	320	23	AB162006	C. pneumoniae BVH
29	31	73.8	323	20	AA135242	Chlamydia pneumoni
30	31	73.8	326	21	AA152545	Arabidopsis thalia
31	31	73.8	335	21	AA131686	Arabidopsis thalia
32	31	73.8	620	21	AA135791	Arabidopsis thalia
33	31	73.8	725	21	AA135790	Arabidopsis thalia
34	31	73.8	735	21	AA131359	Arabidopsis thalia
35	31	73.8	757	21	AA131358	Arabidopsis thalia
36	31	73.8	797	21	AA135789	Arabidopsis thalia
37	31	73.8	823	21	AA131357	Arabidopsis thalia
38	30	71.4	106	22	AA166029	Drosophila melanog
39	30	71.4	140	16	AA166873	Human NF-AT120. H
40	30	71.4	202	21	AA132131	Arabidopsis thalia
41	30	71.4	207	21	AA132130	Arabidopsis thalia
42	30	71.4	238	22	AB127584	Human peptide #235
43	30	71.4	238	22	AB132747	Peptide #253 enco
44	30	71.4	238	22	AB118229	Protein #228 enco
45	30	71.4	238	22	AA153559	Human brain expres

ALIGNMENTS

RESULT 1

ID AAB14276 standard: peptide: 9 AA.

AC AAB14276:

DT 07-MAR-2001 (first entry)

XX

DE HIV-1 p195 epitope.

XX

KW HIV-1: antigenic determinant: vaccine; epitope; p195; AIDS.

XX

OS Human immunodeficiency virus type 1.

XX

PN WO200042068-A2.

PD 20-JUL-2000.

XX

PF 07-JAN-2000; 2000MO-US00372.

XX

PR 11-JAN-1999; 99US-0115430.

PR 06-MAY-1999; 99US-0132760.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Scala G, Chen X, Cohen OJ, Fauci A;

PI WPI; 2000-524189/47.

XX

PT New antigenic peptide is useful as a vaccine for protecting against

PT human immunodeficiency virus-1 -

XX

PS Claim 14; Page 41; 54pp; English.

XX The present peptide is the p195 epitope (antigenic determinant) from  
 CC HIV-1. The present sequence may be used in a vaccine to protect against  
 CC HIV-1 infection. The present peptide was identified by screening random  
 CC peptide libraries using sera from HIV infected subjects who were long  
 CC term non-progressors to AIDS.  
 XX

Sequence 9 AA:  
 SQ

Query Match 95.2%; Score 40; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
 Db 1 KSSGKLISL 9

RESULT 2  
 AAY91252  
 ID AAY91252 standard; peptide: 9 AA.  
 XX  
 AC AAY91252;  
 XX  
 DT 22-MAY-2000 (first entry)  
 XX  
 DE HIV neutralising epitope, SEQ ID NO: 130.  
 XX  
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;  
 KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KM luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KM foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;  
 KM Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;  
 KM cholesterol ester transport protein; anti-arteriosclerotic.  
 KM  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9966957-A2.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US13975.  
 XX  
 PR 20-JUN-1998; 98US-0100412.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 PI Wang CY;  
 XX  
 DR WPI; 2000-160564/14.  
 XX  
 PT New artificial T helper cell epitope and derived immunogens with target  
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis  
 PT or human immune deficiency virus  
 XX  
 PS Claim 13; Page 63; 129pp; English.  
 XX  
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),  
 CC and immunogenic peptides comprising the Th epitopes of the invention  
 CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response.  
 CC Specifically against Plasmodium falciparum, cholesterol ester transport  
 CC protein (CETP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinising hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunosuppression; for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen

CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope  
 CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC immunogens may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th  
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IgE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC Th epitope and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell  
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin  
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.  
 CC  
 XX

Sequence 9 AA:  
 SQ

Query Match 95.2%; Score 40; DB 21; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
 Db 1 KSSGKLISL 9

RESULT 3  
 AAB14273  
 ID AAB14273 standard; peptide: 11 AA.  
 XX  
 AC AAB14273;  
 XX  
 DT 07-MAR-2001 (first entry)  
 XX  
 DE HIV-1 epitope #10.  
 XX  
 XX HIV-1; antigenic determinant; vaccine; epitope; AIDS.  
 KM  
 KW Human immunodeficiency virus type 1.  
 XX  
 OS  
 XX  
 FH Key  
 FH Location/Qualifiers  
 FT Misc-difference 1  
 FT /note= "X is independently an amino acid or sequence of  
 FT amino acids with the proviso that X is not identical to  
 FT the amino acid or amino acids naturally flanking the  
 FT subsequences in HIV-1"  
 FT Misc-difference 11  
 FT /note= "X is independently an amino acid or sequence of  
 FT amino acids with the proviso that X is not identical to  
 FT the amino acid or amino acids naturally flanking the  
 FT subsequences in HIV-1"

PN WO200042068-A2.  
XX 20-JUL-2000.  
XX  
PF 07-JAN-2000; 2000WO-US00372.  
XX  
PR 11-JAN-1999; 99US-0115430.  
PR 06-MAY-1999; 99US-0132760.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Scala G, Chen X, Cohen OJ, Fauci A;  
XX  
DR WPI; 2000-524189/47.  
XX  
PT New antigenic peptide is useful as a vaccine for protecting against  
PT human immunodeficiency virus-1 -  
XX  
PS Claim 1; Page 38; 54pp; English.  
XX  
CC The present peptide is an epitope (antigenic determinant) from  
CC HIV-1. The present sequence may be used in a vaccine to protect against  
CC HIV-1 infection. The present peptide was identified by screening random  
CC peptide libraries using sera from HIV infected subjects who were long  
CC term non-progressors to AIDS.  
XX  
SQ Sequence 11 AA;  
Query Match 95.2%; Score 40; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KSSGKLISL 10  
DB 2 KSSGKLISL 10  
RESULT 4  
AAB0306  
ID AAB0306 standard; Peptide; 11 AA.  
XX  
AC AAB0306;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE HIV-1 gp41/gp120 epitope #5 peptide sequence.  
XX  
KW Fusion protein; antigenic determinant; vaccine;  
KW capsid protein of potato virus X; PVX; CPV; viral hepatitis;  
KW human immunodeficiency virus; HIV; hepatitis C virus; HCV;  
KW Epstein-Barr virus; EBV; influenza virus; antibacterial; anti-HIV;  
KW acquired immunodeficiency syndrome; AIDS; vitucide; hepatotropic.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN EPI167530-A2.  
XX  
PD 02-JAN-2002.  
XX  
PF 11-JUN-2001; 2001EP-0202225.  
XX  
PR 16-JUN-2000; 2000IT-RM00327.  
XX  
PA (CNER ) ENEA ENTE NUOVE TECNOLOGIE ENERGIA.  
PA (SUPE-) INST SUPERIORE DI SANITA'.  
XX  
PI Benvenuto E, Marusic C, Belardelli F, Rizza P, Capone I;  
XX  
DR WPI; 2002-180948/24.  
DR N-PSDB; ABK50059.  
XX  
PT Novel fusion protein, useful for treating e.g. HIV-1 and influenza  
PT virus, comprises an amino terminal antigenic determinant portion of

PT e.g. influenza virus, fused to carboxy terminal portion comprising  
PT capsid protein of potato virus X -  
XX  
PS Claim 4; SEQ ID No 11; 22pp; English.  
XX  
CC The present invention relates to a new fusion protein with an amino  
CC terminal portion comprising an antigenic determinant and a carboxy  
CC terminal portion comprising the capsid protein of potato virus X (PVX)  
CC (CPV) or its variant. The amino terminal portion is fused to the carboxy  
CC terminal portion in such a way that the antigenic determinant is in  
CC frame with CPV or its variant. The molecules of the invention are useful  
CC for preparing a pharmaceutical composition for treating a pathology  
CC associated with bacterial or viral pathogenic agent such as human  
CC immunodeficiency virus (HIV), hepatitis C virus (HCV), Epstein-Barr virus  
CC (EBV) or influenza virus. The medicaments prepared using the molecules of  
CC the invention are useful for treating infectious diseases, preferably  
CC acquired immunodeficiency syndrome (AIDS), viral hepatitis, influenza and  
CC other pathology associated with EBV virus. The present amino acid  
CC sequence represents the HIV gp41/gp120 epitope #5 that was used in the  
CC methods of the invention as a fusion protein, as described above.  
CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied by the  
CC European Patent Office.  
XX  
SQ Sequence 11 AA;  
Query Match 95.2%; Score 40; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KSSGKLISL 10  
DB 2 KSSGKLISL 10  
RESULT 5  
AAB14286  
ID AAB14286 standard; peptide; 20 AA.  
XX  
AC AAB14286;  
XX  
DT 07-MAR-2001 (first entry)  
XX  
DE HIV-1 epitope #4.  
XX  
KW HIV-1; antigenic determinant; vaccine; epitope; AIDS.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO200042068-A2.  
XX  
PD 20-JUL-2000.  
XX  
PF 07-JAN-2000; 2000WO-US00372.  
XX  
PR 11-JAN-1999; 99US-0115430.  
PR 06-MAY-1999; 99US-0132760.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Scala G, Chen X, Cohen OJ, Fauci A;  
XX  
DR WPI; 2000-524189/47.  
XX  
PT New antigenic peptide is useful as a vaccine for protecting against  
PT human immunodeficiency virus-1 -  
XX  
PS Claim 1; Page 38; 54pp; English.  
XX  
CC The present peptide is an epitope (antigenic determinant) from  
CC HIV-1. The present sequence may be used in a vaccine to protect against  
CC HIV-1 infection. The present peptide was identified by screening random  
CC peptide libraries using sera from HIV infected subjects who were long

CC term non-progressors to AIDS.  
 XX  
 SQ Sequence 20 AA;  
 Query Match 95.2%; Score 40; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.33;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLISL 10  
 |||||  
 Db 6 KSSGKLISL 14  
 RESULT 6  
 AAY91258  
 ID AAY91258 standard; peptide: 29 AA.  
 XX AAY91258;  
 AC  
 XX 22-MAY-2000 (first entry)  
 DT  
 XX Modified MWF Th epitope/HIV epitope, SEQ ID NO:136.  
 DE  
 XX Promiscuous T-cell epitope; measles virus F protein; MWF;  
 KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KM luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KM foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;  
 KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
 KM cholesterol ester transport protein; anti-arteriosclerotic.  
 XX  
 OS Chimeric - Measles virus.  
 OS Chimeric - Human immunodeficiency virus type 1.  
 XX  
 PN WO966957-A2.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US13975.  
 XX  
 PR 20-JUN-1998; 98US-0100412.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.  
 XX  
 XX Wang CY;  
 PI  
 XX  
 DR WPI: 2000-160564/14.  
 XX  
 PT New artificial T helper cell epitope and derived immunogens with target  
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis  
 PT or human immune deficiency virus  
 PS  
 PS Claim 13; Page 63; 129pp; English.  
 XX  
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),  
 CC and immunogenic peptides comprising the Th epitopes of the invention  
 CC along with B cell epitopes. The Th epitopes and peptide immunogens  
 CC containing them, are used to induce a T helper cell response,  
 CC specifically against Plasmodium falciparum, cholesterol ester transport  
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,  
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and  
 CC peptide immunogens may be used for prevention and/or treatment of  
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer  
 CC immunotherapy; for inhibition of the action of luteinizing hormone  
 CC releasing hormone (LHRH) for contraception, treatment of hormone-  
 CC dependent cancer, prevention of boar taint in meat, and  
 CC immunocastration); for promoting the growth of animals; or for  
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous  
 CC Th (functional in genetically diverse subjects) into an immunogen  
 CC improves capacity to induce a strong T helper cell-mediated immune  
 CC response, resulting in production of antibodies against a target  
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper  
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope

CC from the measles virus F (MWF) protein and sequences AAY91122-Y91142,  
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the  
 CC MWF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope  
 CC from hepatitis B virus (HBV) surface antigen, and sequences  
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.  
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides  
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197  
 CC is the LHRH target antigenic peptide used in these LHRH antigenic  
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic  
 CC peptides comprising somatostatin and a Th epitope. Somatostatin  
 CC immunogens may be used to promote growth in livestock. AAY91208 is a  
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MWF Th  
 CC epitope/CD4 antigenic peptides which may be used to prevent HIV  
 CC infection of T cells. AAY90212 is a modified version of a human IGE  
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IGE CH3  
 CC antigenic peptides which may be used in the treatment of allergies.  
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)  
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th  
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target  
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MWF Th  
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent  
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a  
 CC CERP peptide and a Th epitope which may be used to prevent or treat  
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257  
 CC are HIV-1 neutralizing B-cell epitopes, and AAY91248-Y91251 and  
 CC AAY91258-Y91273 are antigenic peptides comprising MWF Th and HIV-1 B-cell  
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.  
 CC AAY91198 and AAY91199 are respectively an immunostimulatory lysisin  
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of  
 CC which may optionally be used in the antigenic peptides of the  
 CC invention.  
 CC  
 CC  
 SQ Sequence 29 AA;  
 Query Match 95.2%; Score 40; DB 21; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 0.48;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLISL 10  
 |||||  
 Db 1 KSSGKLISL 9  
 RESULT 7  
 AAY91259  
 ID AAY91259 standard; peptide: 29 AA.  
 XX  
 AC AAY91259;  
 XX  
 DT 22-MAY-2000 (first entry)  
 XX  
 XX Modified MWF Th epitope/HIV epitope, SEQ ID NO:137.  
 DE  
 XX  
 XX Promiscuous T-cell epitope; measles virus F protein; MWF;  
 KM hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;  
 KM luteinizing hormone releasing hormone; LHRH; contraceptive; anticancer;  
 KM somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;  
 KM foot and mouth disease virus; immunoglobulin E; IGE; anti-allergic;  
 KM Plasmodium falciparum; circumsporozoite; antimalarial; CERP;  
 KM cholesterol ester transport protein; anti-arteriosclerotic.  
 XX  
 OS Chimeric - Measles virus.  
 OS Chimeric - Human immunodeficiency virus type 1.  
 XX  
 PN WO966957-A2.  
 XX  
 PD 29-DEC-1999.  
 XX  
 PF 21-JUN-1999; 99WO-US13975.  
 XX  
 PR 20-JUN-1998; 98US-0100412.  
 XX  
 PA (UNBI-) UNITED BIOMEDICAL INC.



ABB36532  
 ID ABB36532 standard; Peptide: 38 AA.  
 AC ABB36532;  
 XX  
 DT 04-FEB-2002 (first entry)  
 XX  
 DE Peptide #4038 encoded by human foetal liver single exon probe.  
 XX  
 KM Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-483447/52.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human fetal liver -  
 PS Claim 27; SEQ ID NO 29167; 639pp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 38 AA;  
 Query Match 78.6%; Score 33; DB 22; Length 38;  
 Best Local Similarity 87.5%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KSSGKLIS 9  
 II:|||||  
 DB 17 KSAAGKLIS 24  
 XX  
 RESULT 10  
 ABB21873  
 ID ABB21873 standard; Protein: 38 AA.  
 AC ABB21873;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Protein #3872 encoded by probe for measuring heart cell gene expression.  
 XX  
 KM Human; gene expression; heart; microarray; vascular system;  
 KM cardiovascular disease; hypertension; cardiac arrhythmia;  
 KM congenital heart disease.  
 XX

OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI: 2001-488899/53.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT hearts -  
 PS Claim 15; SEQ ID NO 23643; 530pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, stratifying, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 38 AA;  
 Query Match 78.6%; Score 33; DB 22; Length 38;  
 Best Local Similarity 87.5%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 2 KSSGKLIS 9  
 II:|||||  
 DB 17 KSAAGKLIS 24  
 XX  
 RESULT 11  
 AAM57294  
 ID AAM57294 standard; Protein: 38 AA.  
 AC AAM57294;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29399.  
 XX  
 KM Human; brain expressed exon; gene expression analysis; probe;  
 KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KM epilepsy; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX



04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0633366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 29399; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 38 AA;

Query Match 78.6%; Score 33; DB 22; Length 38;  
 Best Local Similarity 87.5%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLIS 9  
 ||:|||||  
 Db 17 KSAGKLIS 24

RESULT 12  
 AAM69699  
 ID AAM69699 standard; Protein: 38 AA.  
 XX  
 AC AAM69699;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 30005.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PA WO200157276-A2.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 PF 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0633366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4; SEQ ID NO: 30005; 658bp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 38 AA;

Query Match 78.6%; Score 33; DB 22; Length 38;  
 Best Local Similarity 87.5%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLIS 9  
 ||:|||||  
 Db 17 KSAGKLIS 24

RESULT 13  
 AAM17511  
 ID AAM17511 standard; Protein: 38 AA.

AC AAM17511;  
 DT 12-OCT-2001 (first entry)

DE Peptide #3945 encoded by probe for measuring cervical gene expression.

KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.

OS Homo sapiens.

PA WO200157278-A2.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0633366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.

PF Human genome-derived single exon nucleic acid probes useful for  
 PF analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID NO 22337; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see A110068-A118459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 38 AA;

Query Match 78.6%; Score 33; DB 22; Length 38;  
Best Local Similarity 87.5%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLIS 9  
||:|||||  
DB 17 KSAGKLIS 24

RESULT 14  
AAM30033

ID AAM30033 standard; Protein; 38 AA.

XX AAM30033;

XX 17-OCT-2001 (first entry)

DE Peptide #4070 encoded by probe for measuring placental gene expression.

KM Probe; microarray; human; placenta; antenatal diagnosis;  
genetic disorder.

XX Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for  
analyzing gene expression in human placenta -

PT Claim 27; SEQ ID No 30302; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;  
see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
such probe. The probes are useful for producing a microarray for  
predicting, measuring and displaying gene expression in samples derived  
from human placenta. The probes are useful for antenatal diagnosis of  
human genetic disorders.

XX Sequence 38 AA;

Query Match 78.6%; Score 33; DB 22; Length 38;  
Best Local Similarity 87.5%; Pred. No. 16;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLIS 9  
||:|||||

DB 17 KSAGKLIS 24

RESULT 15  
AAM05183

ID AAM05183 standard; Protein; 38 AA.

XX AAM05183;

XX 09-OCT-2001 (first entry)

DE Peptide #3865 encoded by probe for measuring breast gene expression.

KM Probe; human; breast disease; breast cancer; development disorder;  
inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

PN WO200157270-A2.

PD 09-AUG-2001.

PE 29-JAN-2001; 2001WO-US00661.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression  
in a human breast -

PT Claim 27; SEQ ID No 13923; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes  
(see AAI00010-AAI10067). The present sequence is a peptide encoded by one  
such probe. The probes are useful for measuring human gene expression in  
a human breast sample, where the probe hybridises at high stringency to a  
nucleic acid expressed in the human breast. The probes are useful for  
predicting, diagnosing, grading, staging, monitoring and prognosing  
diseases of the human breast, particularly those diseases with polygenic  
aetiology. The diseases include: breast cancer; disorders of development,  
inflammatory diseases of the breast, fibrocystic changes, proliferative  
breast disease and non-carcinoma tumours.

CC Note: The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 38 AA;

Query Match 78.6%; Score 33; DB 22; Length 38;  
Best Local Similarity 87.5%; Pred. No. 16;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLIS 9  
||:|||||

DB 17 KSAGKLIS 24

Search completed: January 10, 2003, 09:38:22  
Job time : 36 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:15 : Search time 15 Seconds  
(without alignments)  
21.577 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSGKLISLX 11

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	31	73.8	26	2	US-08-394-021-7	Sequence 7, Appl1
2	31	73.8	26	4	US-09-131-551-7	Sequence 7, Appl1
3	30	71.4	140	5	PCT-US94-07297-35	Sequence 35, Appl1
4	30	71.4	250	3	US-09-082-737-3	Sequence 3, Appl1
5	30	71.4	297	4	US-09-068-195-12	Sequence 12, Appl1
6	30	71.4	407	4	US-09-399-913-55	Sequence 55, Appl1
7	30	71.4	414	4	US-09-399-913-55	Sequence 55, Appl1
8	30	71.4	591	5	US-09-082-737-2	Sequence 2, Appl1
9	30	71.4	699	5	PCT-US94-07297-39	Sequence 39, Appl1
10	30	71.4	857	1	US-07-717-331F-2	Sequence 2, Appl1
11	30	71.4	921	1	US-08-396-479B-2	Sequence 2, Appl1
12	30	71.4	921	1	US-08-818-823-2	Sequence 2, Appl1
13	29	69.0	127	4	US-09-134-001C-4589	Sequence 4589, Ap
14	29	69.0	198	2	US-08-825-780-3	Sequence 3, Appl1
15	29	69.0	334	3	US-08-883-526-1	Sequence 1, Appl1
16	29	69.0	334	3	US-08-883-526-4	Sequence 4, Appl1
17	29	69.0	512	4	US-09-302-620B-102	Sequence 102, App
18	28	66.7	14	4	US-09-053-611-14	Sequence 14, Appl
19	28	66.7	71	4	US-09-134-001C-4090	Sequence 4090, App
20	28	66.7	102	1	US-07-778-156-10	Sequence 10, Appl
21	28	66.7	102	2	US-08-422-166-10	Sequence 10, Appl
22	28	66.7	123	4	US-09-329-884-4	Sequence 4, Appl1
23	28	66.7	127	4	US-09-325-932A-176	Sequence 176, App
24	28	66.7	148	4	US-09-325-932A-167	Sequence 167, App
25	28	66.7	169	4	US-09-329-884-12	Sequence 12, Appl
26	28	66.7	169	4	US-09-329-884-14	Sequence 14, Appl
27	28	66.7	175	2	US-08-822-261-1	Sequence 1, Appl1

28	28	66.7	271	4	US-09-318-794A-2	Sequence 2, Appl1
29	28	66.7	271	4	US-09-318-793A-4	Sequence 4, Appl1
30	28	66.7	312	1	US-08-240-783B-6	Sequence 6, Appl1
31	28	66.7	312	1	US-09-084-813-6	Sequence 6, Appl1
32	28	66.7	312	5	PCT-US92-09662-6	Sequence 6, Appl1
33	28	66.7	322	1	US-08-118-270-36	Sequence 36, Appl1
34	28	66.7	322	5	PCT-US93-08528-36	Sequence 36, Appl1
35	28	66.7	329	1	US-08-208-007A-2	Sequence 2, Appl1
36	28	66.7	329	1	US-08-208-007A-7	Sequence 7, Appl1
37	28	66.7	329	1	US-08-330-121B-2	Sequence 2, Appl1
38	28	66.7	329	1	US-08-330-121B-3	Sequence 3, Appl1
39	28	66.7	329	2	US-08-852-807-20	Sequence 20, Appl1
40	28	66.7	329	2	US-08-806-959-2	Sequence 2, Appl1
41	28	66.7	329	3	US-08-964-308-4	Sequence 4, Appl1
42	28	66.7	329	3	US-08-964-308-14	Sequence 14, Appl1
43	28	66.7	329	3	US-08-964-308-15	Sequence 15, Appl1
44	28	66.7	329	3	US-08-964-313-4	Sequence 4, Appl1
45	28	66.7	329	3	US-08-964-313-14	Sequence 14, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-394-021-7  
; Sequence 7, Application US/08394021  
; Patent No. 5830634  
; GENERAL INFORMATION:  
; APPLICANT: BRUST, Stefan  
; APPLICANT: KNAPP, Stefan  
; APPLICANT: GERKEN, Manfred  
; APPLICANT: GUERTLER, Lutz  
; TITLE OF INVENTION: Peptides derived from a retrovirus of  
; TITLE OF INVENTION: the HIV group, and their use  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,021  
; FILING DATE: 23-FEB-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE P 44 05 810.1  
; FILING DATE: 23-FEB-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: SANDERCOCK, Colin G.  
; REGISTRATION NUMBER: 31,298  
; REFERENCE/DOCKET NUMBER: 58315/106/BEAK  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; US-08-394-021-7  
Query Match 73.8%; Score 31; DB 2; Length 26;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

OY 3 SSGKLIS 9  
|||||||  
DB 20 SSGKLIS 26

## RESULT 2

US-09-131-551-7  
Sequence 7, Application US/09131551  
Patent No. 6335158

## GENERAL INFORMATION:

APPLICANT: BRUST, Stefan  
APPLICANT: KNAPP, Stefan  
APPLICANT: GERKEN, Manfred  
APPLICANT: GUERTLER, Lutz  
TITLE OF INVENTION: Peptides derived from a retrovirus of  
TITLE OF INVENTION: the HIV group, and their use  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/131,551  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/394,021  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: SANDERCOCK, Colin G.  
REGISTRATION NUMBER: 31,298  
REFERENCE/DOCKET NUMBER: 58315/106/BEAK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 26 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-131-551-7

Query Match 73.8%; Score 31; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 SSGKLIS 9  
|||||||  
DB 20 SSGKLIS 26

## RESULT 3

PCT-US94-07297-35  
Sequence 35, Application PC/TUS9407297

## GENERAL INFORMATION:

APPLICANT: Arai, Naoko  
APPLICANT: Masuda, Esteban S.  
APPLICANT: Tokumitsu, Hiroshi  
TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN  
TITLE OF INVENTION: TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:

ADDRESSEE: John H. C. Blasdale, Schering-Plough Corporation, W-3-W  
STREET: One Giralda Farms  
CITY: Madison  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07940-1000

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh IIcx  
OPERATING SYSTEM: System Software 7.1  
SOFTWARE: Microsoft Word 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/07297  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/222,626  
FILING DATE: 04-APR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/148,061  
FILING DATE: 05-NOV-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/113,971  
FILING DATE: 30-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/099,998  
FILING DATE: 30-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/088,483  
FILING DATE: 06-JUL-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Blasdale, John H. C.  
REGISTRATION NUMBER: 31,895  
REFERENCE/DOCKET NUMBER: DX0392K4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-822-7398  
TELEFAX: 201-822-7039

## INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:  
LENGTH: 140 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-07297-35

Query Match 71.4%; Score 30; DB 5; Length 140;  
Best Local Similarity 55.6%; Pred. No. 56;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
:||||:|  
DB 53 ESSGRIVSL 61

## RESULT 4

US-09-082-737-3  
Sequence 3, Application US/09082737  
Patent No. 6013500

## GENERAL INFORMATION:

APPLICANT: Minden, Audrey  
TITLE OF INVENTION: PAK4; A No. 6013500el Gene Encoding A Serine/  
TITLE OF INVENTION: Threonine Kinase  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 11230

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/082,737
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 0575/55311
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-082-737-3

Query Match          71.4%; Score 30; DB 3; Length 250;
Best Local Similarity 55.6%; Pred. No. 1e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
DB 18 RSGSKLVAV 26

RESULT 5
US-09-068-195-12
; Sequence 12, Application US/09068195B
; Patent No. 6140078
; GENERAL INFORMATION:
; APPLICANT: Sanders, Jan W.
; APPLICANT: Ledebner, Adrianus M.
; APPLICANT: Venema, Gerard
; APPLICANT: Kok, Jan
; TITLE OF INVENTION: Salt-Inducible Promoter Derivable from a Lactic Acid
; TITLE OF INVENTION: Bacterium, and Its Use in a Lactic Acid Bacterium for
; FILE REFERENCE: Sanders-60113/0252227
; CURRENT APPLICATION NUMBER: US/09/068,195B
; EARLIER FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: PCT/EP97/04755
; EARLIER FILING DATE: 1997-08-20
; EARLIER APPLICATION NUMBER: EP 97200744/7
; EARLIER FILING DATE: 1997-03-13
; EARLIER APPLICATION NUMBER: EP 96202444/4
; EARLIER FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 297
; TYPE: PRT
; ORGANISM: S. gorodnii
;
US-09-068-195-12

Query Match          71.4%; Score 30; DB 4; Length 297;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
DB 5 KSSGKLIRI 13

RESULT 6
US-09-399-913-53
; Sequence 53, Application US/09399913
; Patent No. 6361971
```

```

;
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hui-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 53
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-399-913-53

Query Match          71.4%; Score 30; DB 4; Length 407;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
DB 311 RSGRVSL 319

RESULT 7
US-09-399-913-55
; Sequence 55, Application US/09399913
; Patent No. 6361971
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: Betty, Maria
; APPLICANT: Ling, Hui-Ping
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: POTASSIUM CHANNEL INTERACTORS AND USES THEREFOR
; FILE REFERENCE: MNI-070CP2
; CURRENT APPLICATION NUMBER: US/09/399,913
; CURRENT FILING DATE: 1999-09-21
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; EARLIER APPLICATION NUMBER: USSN 09/350,614
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 09/350,874
; EARLIER FILING DATE: 1999-07-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-399-913-55

Query Match          71.4%; Score 30; DB 4; Length 414;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
```

Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLISL 10  
 :||||:|  
 Db 318 RSSGRVSL 326

RESULT 8  
 US-09-082-737-2  
 : Sequence 2, Application US/09082737  
 : Patent No. 6013500  
 : GENERAL INFORMATION:  
 : APPLICANT: Minden, Audrey  
 : TITLE OF INVENTION: PAK4; A No. 6013500e1 Gene Encoding A Serine/  
 : TITLE OF INVENTION: Threonine Kinase  
 : NUMBER OF SEQUENCES: 12  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Cooper & Dunham LLP  
 : STREET: 1185 Avenue of the Americas  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: U.S.A.  
 : ZIP: 11230  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/09/082,737  
 : FILING DATE:  
 : CLASSIFICATION:  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: White, John P.  
 : REGISTRATION NUMBER: 28,678  
 : REFERENCE/DOCKET NUMBER: 0575/55311  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (212) 278-0400  
 : TELEFAX: (212) 391-0525  
 : INFORMATION FOR SEQ ID NO: 2:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 591 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-09-082-737-2

Query Match 71.4%; Score 30; DB 3; Length 591;  
 Best Local Similarity 55.6%; Pred. NO. 2.4e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLISL 10  
 :||||:|  
 Db 341 RSSGKLAV 349

RESULT 9  
 PCT-US94-07297-39  
 : Sequence 39, Application PC/TUS9407297  
 : GENERAL INFORMATION:  
 : APPLICANT: Arai, Naoko  
 : APPLICANT: Masuda, Esteban S.  
 : APPLICANT: Tokumitsu, Hiroshi  
 : TITLE OF INVENTION: PURIFIED COMPONENTS OF MAMMALIAN  
 : TITLE OF INVENTION: TRANSCRIPTION REGULATION COMPLEXES, AND ANALOGS  
 : NUMBER OF SEQUENCES: 41  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: John H. C. Blasdale, Schering-Plough Corporation, M-3-W  
 : STREET: One Giralda Farms  
 : CITY: Madison  
 : STATE: New Jersey  
 : COUNTRY: USA  
 : ZIP: 07940-1000

COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy disk  
 : COMPUTER: Apple Macintosh IICx  
 : OPERATING SYSTEM: System Software 7.1  
 : SOFTWARE: Microsoft Word 5.1a  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: PCT/US94/07297  
 : FILING DATE:  
 : CLASSIFICATION:  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/222,626  
 : FILING DATE: 04-APR-1994  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/148,061  
 : FILING DATE: 05-NOV-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/113,971  
 : FILING DATE: 30-AUG-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/099,998  
 : FILING DATE: 30-JUL-1993  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: US 08/088,483  
 : FILING DATE: 06-JUL-1993  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Blasdale, John H. C.  
 : REGISTRATION NUMBER: 31,895  
 : REFERENCE/DOCKET NUMBER: DX0392K4  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: 201-822-7398  
 : TELEFAX: 201-822-7039  
 : INFORMATION FOR SEQ ID NO: 39:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 699 amino acids  
 : TYPE: amino acid  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : PCT-US94-07297-39

Query Match 71.4%; Score 30; DB 5; Length 699;  
 Best Local Similarity 55.6%; Pred. NO. 2.8e+02;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 KSSGKLISL 10  
 :||||:|  
 Db 333 ESSGRVSL 341

RESULT 10  
 US-07-717-331F-2  
 : Sequence 2, Application US/07717331F  
 : Patent No. 5484905  
 : GENERAL INFORMATION:  
 : APPLICANT: June Nasrallah, Michael Nasrallah, and Joshua  
 : APPLICANT: Stein  
 : TITLE OF INVENTION: A Receptor Protein Kinase Gene  
 : TITLE OF INVENTION: Encoded At The Self-Incompatibility Locus  
 : NUMBER OF SEQUENCES: 9  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Yahwak & Associates  
 : STREET: 25 Skytop Drive  
 : CITY: Trumbull  
 : STATE: Connecticut  
 : COUNTRY: USA  
 : ZIP: 06611  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy Disk  
 : COMPUTER: Macintosh  
 : OPERATING SYSTEM: MS-DOS  
 : SOFTWARE: Microsoft Word  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/07/717,331F  
 : FILING DATE: June 19th 1991

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: George M. Yahwak  
REGISTRATION NUMBER: 26,824  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203)268-1951  
TELEFAX: (203)268-1951  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 857 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-717-331F-2

Query Match 71.4%; Score 30; DB 1; Length 857;  
Best Local Similarity 75.0%; Pred. No. 3.5e+02;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 SSGKLISL 10  
:|||||  
DB 443 ASGKLISL 450

RESULT 11  
US-08-396-479B-2  
Sequence 2, Application US/08396479B  
Patent No. 5612455  
GENERAL INFORMATION:  
APPLICANT: HOEY, Timothy  
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/396,479B  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59450-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TEL: 210 277299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-396-479B-2

Query Match 71.4%; Score 30; DB 1; Length 921;  
Best Local Similarity 55.6%; Pred. No. 3.7e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
:|||||  
DB 552 ESSGRIVSL 560

RESULT 12  
US-08-818-823-2  
Sequence 2, Application US/08818823  
Patent No. 5708158  
GENERAL INFORMATION:  
APPLICANT: HOEY, Timothy  
TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/818,823  
FILING DATE: 14-MAR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/396,479  
FILING DATE: 02-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Osman, Richard A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: A-59450-1/RAO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 494-8700  
TELEFAX: (415) 494-8771  
TEL: 210 277299  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 921 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-818-823-2

Query Match 71.4%; Score 30; DB 1; Length 921;  
Best Local Similarity 55.6%; Pred. No. 3.7e+02;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
:|||||  
DB 552 ESSGRIVSL 560

RESULT 13  
US-09-134-001C-4589  
Sequence 4589, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC  
FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
PRIOR FILING DATE: 1996-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 4589  
LENGTH: 127  
TYPE: PRT  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4589

Query Match 69.0%; Score 29; DB 4; Length 127;  
 Best Local Similarity 75.0%; Pred. No. 81;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLIS 9  
 11111111  
 Db 17 KSSGKLIS 24

RESULT 14

US-08-825-780-3  
 ; Sequence 3, Application US/08825780  
 ; Patent No. 5834238  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Shah, Purvi  
 ; TITLE OF INVENTION: NOVEL HUMAN GTP BINDING P  
 ; TITLE OF INVENTION: ROTEIN  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/825,780  
 ; FILING DATE: Filed Herewith  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0264 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 198 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 436564  
 ; US-08-825-780-3

Query Match 69.0%; Score 29; DB 2; Length 198;  
 Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
 11111111  
 Db 23 KSSGKLISL 31

RESULT 15  
 US-08-883-526-1  
 ; Sequence 1, Application US/08883526  
 ; Patent No. 6033893  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Guegler, Karl J.

APPLICANT: Corley, Neil C.  
 APPLICANT: Shah, Purvi  
 TITLE OF INVENTION: NEW HUMAN CATHEPSIN  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/883,526  
 FILING DATE: Herewith  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy J.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/DOCKET NUMBER: PF-0331 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-845-4166  
 TELEX:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 334 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: THYMNOTO2  
 ; CLONE: 347021,  
 ; US-08-883-526-1

Query Match 69.0%; Score 29; DB 3; Length 334;  
 Best Local Similarity 55.6%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
 11111111  
 Db 153 KSSGKLISL 161

Search completed: January 10, 2003, 09:39:37  
 Job time : 16 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:38:44 ; Search time 11 Seconds  
(without alignments)  
19.401 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSSGKLISLX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCR\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCRUS\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	78.6	38	10	US-09-864-761-37171
2	31	73.8	26	12	US-10-000-321-7
3	31	73.8	93	10	US-09-925-297-749
4	30	71.4	238	9	US-09-864-761-33527
5	30	71.4	362	9	US-09-870-759-80
6	30	71.4	364	10	US-09-864-761-33528
7	30	71.4	407	10	US-09-350-874-53
8	30	71.4	414	10	US-09-350-874-55
9	30	71.4	927	10	US-09-841-786-3
10	30	71.4	3241	10	US-09-841-786-1
11	29	69.0	69	10	US-09-864-761-34136
12	29	69.0	201	10	US-09-925-300-1357
13	29	69.0	313	10	US-09-755-017-2
14	29	69.0	334	9	US-10-028-072-12
15	29	69.0	334	10	US-09-529-063-1
16	29	69.0	779	10	US-09-749-601A-12
17	29	69.0	1189	9	US-09-738-626-4140
18	28	66.7	123	10	US-09-865-159-4
19	28	66.7	169	10	US-09-865-159-12

20	28	66.7	169	10	US-09-865-159-14	Sequence 14, Appl
21	28	66.7	175	9	US-09-978-295A-452	Sequence 452, App
22	28	66.7	175	9	US-09-978-697-452	Sequence 452, App
23	28	66.7	175	9	US-09-978-192A-452	Sequence 452, App
24	28	66.7	175	9	US-09-999-832A-452	Sequence 452, App
25	28	66.7	175	9	US-09-978-189-452	Sequence 452, App
26	28	66.7	175	12	US-10-052-586-424	Sequence 424, App
27	28	66.7	200	9	US-09-738-626-698	Sequence 6458, App
28	28	66.7	271	9	US-09-738-626-3633	Sequence 3633, App
29	28	66.7	329	10	US-09-953-956-2	Sequence 2, Appl1
30	28	66.7	329	12	US-10-114-464-2	Sequence 2, Appl1
31	28	66.7	329	12	US-10-114-464-7	Sequence 7, Appl1
32	28	66.7	329	12	US-10-114-464-7	Sequence 37, Appl1
33	28	66.7	356	9	US-09-771-382-37	Sequence 1693, Ap
34	28	66.7	361	10	US-09-925-300-1693	Sequence 38, Appl
35	28	66.7	382	9	US-09-771-382-38	Sequence 25, Appl
36	28	66.7	407	9	US-09-771-382-25	Sequence 25, Appl
37	28	66.7	433	9	US-09-771-382-26	Sequence 35, Appl
38	28	66.7	461	9	US-09-771-382-35	Sequence 36, Appl
39	28	66.7	462	9	US-09-771-382-36	Sequence 27, Appl
40	28	66.7	502	9	US-09-771-382-27	Sequence 23, Appl
41	28	66.7	512	9	US-09-771-382-23	Sequence 24, Appl
42	28	66.7	513	9	US-09-771-382-24	Sequence 33, Appl
43	28	66.7	540	9	US-09-771-382-33	Sequence 34, Appl
44	28	66.7	541	9	US-09-771-382-34	Sequence 3, Appl1
45	28	66.7	589	9	US-09-771-382-3	

#### ALIGNMENTS

RESULT 1  
US-09-864-761-37171  
Sequence 37171, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT FILING DATE: 2001-05-23  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30

```

: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 37171
: LENGTH: 38
: TYPE: prt
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC009296.1
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.86
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
: OTHER INFORMATION: EXPRESSED IN FBEL100, SIGNAL = 0.85
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.77
: OTHER INFORMATION: EST_HUMAN HIT: BE883574.1, EVALUATE 1.10e+00
: OTHER INFORMATION: SWISSPROT HIT: P16415, EVALUATE 8.20e-02
: US-09-864-761-37171

```

```

Query Match      78.8%; Score 31; DB 10; Length 38;
Best Local Similarity 87.5%; Pred. No. 1.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2 KSSGKLIS 9
   11:11111
Db 17 KSAGKLIS 24

```

```

RESULT 2
: US-10-000-321-7
: Sequence 7, Application US/10000321
: Patent No. US20020123039A1
: GENERAL INFORMATION:
: APPLICANT: BRUST, Stefan
: KNAPP, Stefan
: GERKEN, Manfred
: GUERTLER, Lutz
: TITLE OF INVENTION: Peptides derived from a retrovirus of
: the HIV group, and their use
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/000,321
: FILING DATE: 04-Dec-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 09/131,551
: FILING DATE: <Unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: SANDERCOCK, Colin G.
: REGISTRATION NUMBER: 31,298

```

```

: REFERENCE/DOCKET NUMBER: 58315/106/BEAK
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELE: 904136
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 26 amino acids
: TYPE: amino acid
: STRANDEDNESS: <Unknown>
: TOPOLOGY: linear
: SEQUENCE DESCRIPTION: SEQ ID NO: 7:
: US-10-000-321-7

```

```

Query Match      73.8%; Score 31; DB 12; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3 SSGKLIS 9
   1111111
Db 20 SSGKLIS 26

```

```

RESULT 3
: US-09-925-297-749
: Sequence 749, Application US/09925297
: Patent No. US20020081659A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
: FILE REFERENCE: PA105
: CURRENT APPLICATION NUMBER: US/09/925,297
: CURRENT FILING DATE: 2001-08-10
: PRIOR APPLICATION NUMBER: PCT/US00/05989
: PRIOR FILING DATE: 2000-03-08
: PRIOR APPLICATION NUMBER: 60/124,270
: PRIOR FILING DATE: 1999-03-12
: NUMBER OF SEQ ID NOS: 928
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 749
: LENGTH: 93
: TYPE: prt
: ORGANISM: Homo sapiens
: US-09-925-297-749

```

```

Query Match      73.8%; Score 31; DB 10; Length 93;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 KSSGKLISL 10
   111111:
Db 63 KLSGKLVSII 71

```

```

RESULT 4
: US-09-864-761-33527
: Sequence 33527, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366

```

;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00663  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00662  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 33527  
;; LENGTH: 238  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; OTHER INFORMATION: MAP TO AL050331.11  
;; OTHER INFORMATION: EXPRESSED IN BR474, SIGNAL = 4  
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.4  
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5  
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6  
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4  
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2  
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
;; OTHER INFORMATION: EST\_HUMAN HIT: BF346321.1, EVALUATE 5.00e-62  
;; OTHER INFORMATION: EST\_HUMAN HIT: AM390205.1, EVALUATE 3.00e-60  
;; OTHER INFORMATION: SWISSPROT HIT: Q01105, EVALUATE 3.00e-39  
;; US-09-864-761-33527

Query Match 71.4%; Score 30; DB 10; Length 238;  
Best Local Similarity 55.6%; Pred. No. 49;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KSSGRLISL 10  
:||||:||||  
Db 155 RSSGRVSL 163

RESULT 5  
US-09-870-759-80  
; Sequence 80, Application US/09870759  
; Patent No. US2002017751A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14

;; PRIOR APPLICATION NUMBER: US 60/208,128  
;; PRIOR FILING DATE: 2000-05-30  
;; NUMBER OF SEQ ID NOS: 166  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 80  
;; LENGTH: 362  
;; TYPE: PRT  
;; ORGANISM: Bacteriophage T270  
;; US-09-870-759-80

Query Match 71.4%; Score 30; DB 9; Length 362;  
Best Local Similarity 77.8%; Pred. No. 78;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KSSGRLISL 10  
:||||:||||  
Db 239 KSSDRLISL 247

RESULT 6  
US-09-864-761-33528  
; Sequence 33528, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Aeomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/632,366  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: GB 24263.6  
;; PRIOR FILING DATE: 2000-10-04  
;; PRIOR APPLICATION NUMBER: US 60/236,359  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: PCT/US01/00666  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00667  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00664  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00669  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00665  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00668  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00661  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: PCT/US01/00670  
;; PRIOR FILING DATE: 2001-01-30  
;; PRIOR APPLICATION NUMBER: US 60/234,687  
;; PRIOR FILING DATE: 2000-09-21  
;; PRIOR APPLICATION NUMBER: US 09/608,408  
;; PRIOR FILING DATE: 2000-06-30  
;; PRIOR APPLICATION NUMBER: US 09/774,203  
;; PRIOR FILING DATE: 2001-01-29  
;; NUMBER OF SEQ ID NOS: 49117  
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
;; SEQ ID NO 33528  
;; LENGTH: 364

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050331.11
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.9
; OTHER INFORMATION: SWISSPROT HIT: Q01534, EVALUE 3.00e-37
; OTHER INFORMATION: EST_HUMAN HIT: AW937326.1, EVALUE 3.00e-76
US-09-864-761-33528

Query Match          71.4%; Score 30; DB 10; Length 364;
Best Local Similarity 55.6%; Pred. No. 79;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 326 RSSGRVSL 334

RESULT 7
US-09-350-874-53
; Sequence 53, Application US/09350874
; Patent No. US20020019020A1
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: METHODS FOR TREATING CARDIOVASCULAR DISORDERS
; FILE REFERENCE: NMT-069
; CURRENT APPLICATION NUMBER: US/09/350,874
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-350-874-53

Query Match          71.4%; Score 30; DB 10; Length 407;
Best Local Similarity 55.6%; Pred. No. 89;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 311 RSSGRVSL 319

RESULT 8
US-09-350-874-55
; Sequence 55, Application US/09350874
; Patent No. US20020019020A1
; GENERAL INFORMATION:
; APPLICANT: Rhodes, Kenneth
; APPLICANT: An, Wenqian
; TITLE OF INVENTION: METHODS FOR TREATING CARDIOVASCULAR DISORDERS
; FILE REFERENCE: NMT-069
; CURRENT APPLICATION NUMBER: US/09/350,874
```

```
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: USSN 60/110,277
; EARLIER FILING DATE: 1998-11-30
; EARLIER APPLICATION NUMBER: USSN 60/110,033
; EARLIER FILING DATE: 1998-11-25
; EARLIER APPLICATION NUMBER: USSN 60/109,333
; EARLIER FILING DATE: 1998-11-20
; EARLIER APPLICATION NUMBER: USSN 09/298,731
; EARLIER FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-874-55

Query Match          71.4%; Score 30; DB 10; Length 414;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 318 RSSGRVSL 326

RESULT 9
US-09-841-786-3
; Sequence 3, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 927
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-3

Query Match          71.4%; Score 30; DB 10; Length 927;
Best Local Similarity 55.6%; Pred. No. 2.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 129 KANGKLISI 137

RESULT 10
US-09-841-786-1
; Sequence 1, Application US/09841786
; Patent No. US20020054883A1
; GENERAL INFORMATION:
; APPLICANT: NAGARAJA, T. G.
; APPLICANT: STEWART, GEORGE C.
; APPLICANT: NARAYANAN, SANJEEV K.
; APPLICANT: CHENGAPPA, M. M.
; TITLE OF INVENTION: RECOMBINANT FUSOBACTERIUM NECROPHORUM LEUKOTOXIN
; TITLE OF INVENTION: VACCINE AND PREPARATION THEREOF
; FILE REFERENCE: 30296
; CURRENT APPLICATION NUMBER: US/09/841,786
; CURRENT FILING DATE: 2001-04-24
```

```
; PRIOR APPLICATION NUMBER: 09/558,257
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3241
; TYPE: PRT
; ORGANISM: Fusobacterium necrophorum
US-09-841-786-1

Query Match
Best Local Similarity 71.4%; Score 30; DB 10; Length 3241;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
Db 435 KANGKLISI 443

RESULT 11
US-09-864-761-34136
; Sequence 34136, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34136
; LENGTH: 69
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000118.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.5
; OTHER INFORMATION: SWISSPROT HIT: P18886, EVALU8 8.20e-01
; OTHER INFORMATION: EST_HUMAN HIT: A0145181.1, EVALU8 5.50e+00
US-09-864-761-34136

Query Match
Best Local Similarity 69.0%; Score 29; DB 10; Length 69;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 3 SSGKLISL 10
Db 18 SQGKLISL 25

RESULT 12
US-09-925-300-1357
; Sequence 1357, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05968
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1357
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1357

Query Match
Best Local Similarity 69.0%; Score 29; DB 10; Length 201;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
Db 26 KKSGLIVEL 34

RESULT 13
US-09-755-017-2
; Sequence 2, Application US/09755017
; Patent No. US20010034438A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20010034438A1el Human Membrane Proteins and
; TITLE OF INVENTION: Polynucleotides Encoding the Same
```

```

; FILE REFERENCE: LEX-0115-USA
; CURRENT APPLICATION NUMBER: US/09/755,017
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/175,764
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 313
; TYPE: PRF
; ORGANISM: Homo Sapien
US-09-755-017-2

Query Match      69.0%; Score 29; DB 10; Length 313;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY      2 KSSGRLSL 10
Db      268 KDOGKMWSL 276

RESULT 14
US-10-028-072-12
; Sequence 12, Application US/10038072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285

```

```

; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27

```

PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086414  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/086430  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088730  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088741  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090445  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090538  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090863  
PRIOR FILING DATE: 1998-06-26  
PRIOR APPLICATION NUMBER: 60/091360

PRIOR FILING DATE: 1998-07-01  
PRIOR APPLICATION NUMBER: 60/091519  
PRIOR FILING DATE: 1998-07-02  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07

Query Match 69.0%; Score 29; DB 9; Length 334;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
: : : : :  
Db 153 RKTGKLVS 161

RESULT 15  
US-09-529-063-1  
Sequence 1, Application US/09529063  
Patent No. US20020102542A1  
GENERAL INFORMATION:  
APPLICANT: FUKUSHIMA, DAIKICHI  
APPLICANT: SHIBAYAMA, SHIRO  
TADA, HIDEAKI  
TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF  
TITLE OF INVENTION: THE BOTH  
FILE REFERENCE: 058769  
CURRENT APPLICATION NUMBER: US/09/529.063  
CURRENT FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: PCT/JP98/04514  
PRIOR FILING DATE: 1998-10-06  
PRIOR APPLICATION NUMBER: JP 9-274674  
PRIOR FILING DATE: 1997-10-07  
NUMBER OF SEQ ID NOS: 117  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 334  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-529-063-1

Query Match 69.0%; Score 29; DB 10; Length 334;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
: : : : :  
Db 153 RKTGKLVS 161

Search completed: January 10, 2003, 09:46:43  
Job time : 12 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:14 : Search time 15 seconds  
(without alignments)  
70.499 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSSGKLISLX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	88.1	763	2 F96693	hypothetical prote
2	35	83.3	809	2 S33533	heat shock protein
3	34	81.0	88	2 G96907	hypothetical prote
4	34	81.0	206	1 A45079	photosynthetic rea
5	34	81.0	1465	2 T23056	hypothetical prote
6	33	78.6	559	2 T09038	NADH dehydrogenase
7	33	78.6	659	2 A75029	h+-transporting AT
8	32	76.2	325	2 I40787	replication protei
9	32	76.2	325	2 I40788	replication protei
10	32	76.2	325	2 I40829	replication protei
11	32	76.2	325	2 A32310	replication protei
12	32	76.2	325	2 T00233	replication protei
13	32	76.2	856	1 UC2482	S-receptor kinase
14	31	73.8	218	2 F72214	hypothetical prote
15	31	73.8	286	2 T16241	hypothetical prote
16	31	73.8	292	2 JC4741	mosquitocidal toxi
17	31	73.8	312	2 T17118	protein kinase cdc
18	31	73.8	320	2 C86568	CTD50 hypothetical
19	31	73.8	320	2 B72057	conserved hypothet
20	31	73.8	412	2 G64059	probable serine tr
21	31	73.8	560	2 I50372	hypothetical prote
22	31	73.8	672	2 S75001	hypothetical prote
23	31	73.8	817	2 T37989	hypothetical prote
24	31	73.8	817	2 S39558	DNA mismatch repai
25	31	73.8	823	2 T09882	HSP90 homolog - Ma
26	31	73.8	852	2 T00994	heat shock protein
27	31	73.8	861	2 B70866	hypothetical prote
28	31	73.8	1131	2 T15617	hypothetical prote
29	31	73.8	1253	2 T40302	hypothetical prote

30	31	73.8	1388	2 T17269	hypothetical prote
31	31	73.8	1711	1 A47392	chromodomain-helic
32	31	73.8	4574	2 G02520	plectin - human
33	31	73.8	4684	2 A59404	plectin [imported]
34	31	73.8	4687	1 A39638	plectin - rat
35	30	71.4	34	2 F70242	hypothetical prote
36	30	71.4	136	2 B64754	glucose 1-dehydrog
37	30	71.4	248	2 AC1657	hypothetical prote
38	30	71.4	249	2 A99667	hypothetical prote
39	30	71.4	297	2 A41898	positive regulator
40	30	71.4	299	2 E85517	probable lysR-like
41	30	71.4	311	2 C84807	probable cell divi
42	30	71.4	356	2 G96806	thiamatin-like pro
43	30	71.4	359	2 S63649	probable gene Arpa
44	30	71.4	363	2 T05624	cinnamyl-alcohol d
45	30	71.4	363	2 T06726	cysteine proteinase

## ALIGNMENTS

RESULT 1  
F96693  
hypothetical protein F1019.5 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: F96693  
R:Theologis, A.; Eckert, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Hulzar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Kowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: AB6141; MUID:21016719; PMID:11130712  
A:Accession: F96693  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-763 <STO>  
A:Cross-references: GB:AE005173; NID:g9755448; PIDN:AAF98209.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F1019.5  
A:Map position: 1  
Query Match  
Best Local Similarity 88.1% Score 37; DB 2; Length 763;  
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 XKSSGKLISL 10  
Db 477 KNSGKLISL 485  
RESULT 2  
S33533  
heat shock protein 90 homolog precursor - barley  
N:Alternate names: GRP94 protein homolog  
C:Species: Hordeum vulgare (barley)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 20-Aug-1999  
R:Walter-Larsen, H.; Brandt, J.; Collinge, D.B.; Thordal-Christensen, H.  
Plant Mol. Biol. 21, 1097-1108, 1993  
A:Title: A pathogen-induced gene of barley encodes a HSP90 homologue showing striking  
A:Reference number: S33533; MUID:93257625; PMID:8490130  
A:Accession: S33533  
A:Molecule type: mRNA  
A:Residues: 1-809 <WAL>  
A:Cross-references: EMBL:X67960; NID:g22651; PIDN:CAA48143.1; PID:g22652  
C:Superfamily: heat shock protein 90  
C:Keywords: glycoprotein

F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-809/Product: heat shock protein 90 homolog #status predicted <MAT>  
F:111,410,450,617/Binding site: carbohydrate (Asn) (covalent) #status predicted

## Query Match

83.3%: Score 35; DB 2; Length 809;  
Best Local Similarity 77.8%; Pred. No. 26;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
|||:|:|:  
DB 543 KSDGKLIVSL 551

## RESULT 3

G96907

hypothetical protein CAC0064 [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: G96907

R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium CLO

A:Reference number: A86900; MUID:21359325; PMID:21359325

A:Accession: G96907

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-88 &lt;KUR&gt;

A:Cross-references: CLO:AE001437; PIDN:AAK78050.1; PID:G15022886; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC624

C:Genetics:

A:Gene: CAC0064

## Query Match

81.0%: Score 34; DB 2; Length 88;  
Best Local Similarity 77.8%; Pred. No. 4.6;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
|||:|:|:  
DB 17 KSSGELISL 25

## RESULT 4

A45079

photosynthetic reaction center cytochrome c551 - Chlorobium vibrioforme

N:Alternate names: photosynthetic reaction center complex 18k protein

C:Species: Chlorobium vibrioforme

C:Date: 10-Jun-1993 #sequence\_revision 02-Jul-1996 #text\_change 03-Mar-2000

C:Accession: A45079

R:Okels, J.S.; Kjaer, B.; Hansson, O.; Svendsen, I.; Moller, B.L.; Scheller, H.V.

J. Biol. Chem. 267, 21139-21145, 1992

A:Title: A membrane-bound monoheme cytochrome c551 of a novel type is the immediate elec

A:Reference number: A45079; MUID:93016035; PMID:1383218

A:Accession: A45079

A:Molecule type: DNA; protein

A:Residues: 1-206 &lt;OKK&gt;

A:Cross-references: GB:M95751; NID:G144472; PIDN:AAA23110.1; PID:G144473

A:Experimental source: f. thiosulfatophilum 8327

A:Note: sequence extracted from NCBI backbone (NCBIN:116201, NCBI:P.116202)

A:Note: part of this sequence, including the amino end of the mature protein, confirmed

A:Note: authors predicted the Met axial ligand based on its predicted location on the pe

C:Genetics:

A:Gene: cyca

C:Complex: tightly bound to the photosynthetic reaction center complex

A:Description: electron donor to photooxidized P840 of the photosynthetic reaction cente

C:Superfamily: Chlorobium photosynthetic reaction center cytochrome c551

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; photosynthesis

F:1-9/Product: photosynthetic reaction center cytochrome c551 #status experimental &lt;MA

F:10-32/Domain: intracellular #status predicted &lt;INT&gt;

F:33-48/Domain: periplasmic #status predicted &lt;PER1&gt;

F:49-67/Domain: transmembrane #status predicted &lt;TM2&gt;

F:78-95/Domain: transmembrane #status predicted <TM3>  
F:96-206/Domain: periplasmic #status predicted <PER2>  
F:156-155/Binding site: heme (Cys) (covalent) #status predicted  
F:156-182/Binding site: heme iron (His, Met) (axial ligands) #status predicted

## Query Match

81.0%: Score 34; DB 1; Length 206;  
Best Local Similarity 77.8%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
|||:|:|:  
DB 4 KSSGKLIAL 12

## RESULT 5

T23056

hypothetical protein H06001.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000

C:Accession: T23056

R:Barlow, K.

submitted to the EMBL Data Library, March 1997

A:Reference number: 219665

A:Accession: T23056

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1465 &lt;NLL&gt;

A:Cross-references: EMBL:T292970; PIDN:CA807481.1; GSPDB:GN00019; CESP:H06001.2

A:Experimental source: clone H06001

C:Genetics:

A:Gene: CESP:H06001.2

A:Map position: 1

A:Introns: 44/1; 91/3; 170/3; 377/3; 494/3; 1046/2; 1099/3; 1242/3; 1298/1

C:Superfamily: CHD-1 protein; chromobox homology

## Query Match

81.0%: Score 34; DB 2; Length 1465;  
Best Local Similarity 88.9%; Pred. No. 78;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
|||:|:|:  
DB 708 KSSGKLIAL 716

## RESULT 6

T09038

NADH dehydrogenase (ubiquinone) chain ND11 homolog F26K10.100 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2000

C:Accession: T09038

R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Ban

submitted to the Protein Sequence Database, June 1999

A:Reference number: 216533

A:Accession: T09038

A:Molecule type: DNA

A:Residues: 1-559 &lt;BEV&gt;

A:Cross-references: EMBL:AL049803; ATSP:F26K10.100; GSPDB:GN00062

A:Experimental source: cultivar Columbia; BAC clone F26K10

C:Genetics:

A:Gene: ATSP:F26K10.100

A:Map position: 4

A:Introns: 27/2; 113/3; 182/3; 266/2; 374/3; 462/3; 503/2; 552/2

C:Superfamily: NADH dehydrogenase

## Query Match

78.6%: Score 33; DB 2; Length 559;  
Best Local Similarity 66.7%; Pred. No. 49;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
|||:|:|:  
DB 305 KSSGELVSI 313

```
RESULT 7
A:75029
h+-transporting ATP synthase, chain I (atpI) PAB1180 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: A75029
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: A75029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659 <KAW>
A:Cross-references: GB:AJ248288; GB:AL096836; NID:g5458960; PIDN:CAB50671.1; PID:el51657
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1180

Query Match
Best Local Similarity 77.8%; Score 33; DB 2; Length 659;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 206 KDSGKIVTL 214

RESULT 8
A:140787
replication protein repI - Escherichia coli plasmid ColV3-K30
C:Species: Escherichia coli
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 26-Aug-1999
C:Accession: I40787; A32842
R:Gibbs, M.D.; Spiers, A.J.; Bergquist, P.L.
Plasmid 29, 165-179, 1993
A:Title: RepFIB: a basic replicon of large plasmids.
A:Reference number: I40787; MUID:93361574; PMID:8356112
A:Accession: I40787
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: GB:L01251; NID:g144671; PIDN:AAA71881.1; PID:g144672
A:Experimental source: plasmid ColV3-K30
R:Perez-Casal, J.F.; Gamme, A.E.; Crosa, J.H.
J. Bacteriol. 171, 2195-2201, 1989
A:Title: Nucleotide sequence analysis and expression of the minimum REPI replication reg
A:Reference number: A32842; MUID:89197795; PMID:2703470
A:Accession: A32842
A:Molecule type: DNA
A:Residues: 25-325 <PER>
A:Cross-references: GB:M24908; NID:g144699; PIDN:AAA23197.1; PID:g144700
C:Genetics:
A:Gene: repA
A:Genome: plasmid
A:Start codon: GTG
C:Superfamily: repA protein

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 325;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 3 KSSGELVTL 11

RESULT 9
A:140788
replication protein A - Escherichia coli plasmid ColV
C:Species: Escherichia coli
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 26-Aug-1999
C:Accession: I40788
```

```
R:Gibbs, M.D.; Spiers, A.J.; Bergquist, P.L.
Plasmid 29, 165-179, 1993
A:Title: RepFIB: a basic replicon of large plasmids.
A:Reference number: I40787; MUID:93361574; PMID:8356112
A:Accession: I40788
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: GB:L01250; NID:g144697; PIDN:AAA71880.1; PID:g144698
A:Experimental source: Plasmid ColV
C:Genetics:
A:Gene: repA
A:Genome: plasmid
A:Start codon: GTG
C:Superfamily: repA protein

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 325;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 3 KSSGELVTL 11

RESULT 10
A:140829
replication protein A - Escherichia coli plasmid ColVbtrp
C:Species: Escherichia coli
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 26-Aug-1999
C:Accession: I40829
R:Gibbs, M.D.; Spiers, A.J.; Bergquist, P.L.
Plasmid 29, 165-179, 1993
A:Title: RepFIB: a basic replicon of large plasmids.
A:Reference number: I40787; MUID:93361574; PMID:8356112
A:Accession: I40829
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: GB:L01256; NID:g144954; PIDN:AAA71886.1; PID:g144955
A:Experimental source: plasmid ColVbtrp
C:Genetics:
A:Gene: repA
A:Genome: plasmid
A:Start codon: GTG
C:Superfamily: repA protein

Query Match
Best Local Similarity 76.2%; Score 32; DB 2; Length 325;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10
Db 3 KSSGELVTL 11

RESULT 11
A:32310
replication protein RepFIB (P307 replicon) - Escherichia coli plasmid IncF
C:Species: Escherichia coli
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 21-Aug-1998
C:Accession: A32310
R:Saut, D.; Spiers, A.J.; McNulty, J.; Gibbs, M.G.; Bergquist, P.L.; Hill, D.F.
J. Bacteriol. 171, 2697-2707, 1989
A:Title: Nucleotide sequence and replication characteristics of RepFIB, a basic repli
A:Reference number: A32310; MUID:89213960; PMID:2651415
A:Accession: A32310
A:Molecule type: DNA
A:Residues: 1-325 <SAU>
C:Genetics:
A:Gene: repA
A:Genome: plasmid
A:Start codon: GTG
C:Superfamily: repA protein
C:Keywords: plasmid copy control
```

Query Match 76.2%; Score 32; DB 2; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKILSL 10  
 |||||:|:|  
 Db 3 KSSGELVTL 11

# RESULT 12

T00233  
 replication protein RepFIB - Escherichia coli plasmid p0157

C:Species: Escherichia coli  
 C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 20-Jun-2000  
 C:Accession: T00233  
 R:Markino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shingawa, H.  
 DNA Res. 5, 1-9, 1998  
 A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic E. coli O157:H7, MUD:98290540; PMID:9628576  
 A:Accession: T00233  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-325 <MAK>  
 A:Cross-references: EMBL:AB011549; PIDN:BAA31780.1  
 A:Experimental source: strain EHCC O157:H7, substrain R1MD 0509952  
 C:Genetics:  
 A:Gene: repFIB  
 A:Genome: plasmid p0157  
 C:Superfamily: repA protein

Query Match 76.2%; Score 32; DB 2; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 47;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKILSL 10  
 |||||:|:|  
 Db 3 KSSGELVTL 11

# RESULT 13

JC2482  
 S-receptor kinase (EC 2.7.1.-) 12 precursor - field mustard (fragment)

N:Alternate names: receptor protein kinase K, SRK12  
 C:Species: Brassica campestris (field mustard)  
 C:Date: 22-Apr-1995 #sequence\_revision 01-Aug-1997 #text\_change 26-Feb-1999  
 C:Accession: JC2482  
 R:Yamakawa, S.; Watanabe, M.; Hinata, K.; Suzuki, A.; Isogai, A.  
 Biosci. Biotechnol. Biochem. 59, 161-162, 1995  
 A:Title: The sequences of S-receptor kinase (SRK) involved in self-incompatibility and self-incompatibility in field mustard  
 A:Reference number: JC2481; MUID:95201375; PMID:7765971  
 A:Accession: JC2482

A:Molecule type: mRNA  
 A:Residues: 1-856 <YAM>  
 A:Cross-references: DDBJ:D38564  
 C:Genetics:  
 A:Gene: SRK  
 A:Note: locus is highly polymorphic

C:Function:  
 A:Description: involved in preventing fertilization between plants having the same S-locus  
 C:Superfamily: S-receptor kinase; protein kinase homology; S-locus-specific glycoprotein  
 C:Keywords: ATP; glycoprotein; magnesium; phosphotransferase; receptor; serine/threonine  
 F:1-31/Domain: signal sequence (fragment) #status predicted <SIG>  
 F:32-856/Product: S-receptor kinase 12 #status predicted <MAK>  
 F:42-436/Domain: S-locus-specific glycoprotein homology <SSG>  
 F:447-466/Domain: transmembrane #status predicted <TMW>  
 F:526-812/Domain: protein kinase homology <KIN>  
 F:534-542/Region: protein kinase ATP-binding motif  
 F:48-79, 123, 217, 247, 317, 392/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:555, 571, 652, 654/Active site: Lys, Glu, Asp, Lys #status predicted  
 F:657, 661/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 76.2%; Score 32; DB 1; Length 856;  
 Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKILSL 10  
 |||||:|:|  
 Db 442 KANGKILSL 450

# RESULT 14

F72214  
 hypothetical protein - Thermotoga maritima (strain MSB)

C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: F72214  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome Nucleotide sequencing  
 A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: F72214  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-218 <ARN>  
 A:Cross-references: GB:AE001814; GB:AE000512; NID:g4982332; PIDN:AA036822.1; PID:g498  
 A:Experimental source: strain MSB  
 C:Genetics:  
 A:Gene: TM1757

Query Match 73.8%; Score 31; DB 2; Length 218;  
 Best Local Similarity 77.8%; Pred. No. 52;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKILSL 10  
 |||||:|:|  
 Db 25 KSSGKIVSL 33

# RESULT 15

T16241  
 hypothetical protein F33G12.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T16241  
 R:Nhan, M.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid F33G12.  
 A:Reference number: Z18484  
 A:Accession: T16241

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-286 <NHA>  
 A:Cross-references: EMBL:U01278; NID:g1086899; PID:g1086900; PIDN:AAA82480.1; CESP:F3  
 C:Genetics:  
 A:Gene: CESP:F33G12.2  
 A:introns: 26/1; 82/1; 156/3; 200/2; 235/2

Query Match 73.8%; Score 31; DB 2; Length 286;  
 Best Local Similarity 75.0%; Pred. No. 69;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKILSL 9  
 |||||:|:|  
 Db 192 KSSGKILSL 199

Search completed: January 10, 2003, 09:37:40  
 Job time : 16 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:14 : Search time 11 Seconds

(without alignments)  
41.476 Million cell updates/sec

Title: US-09-869-003-1

Perfect score: 42

Sequence: 1 XKSSGKLSLX 11

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	35	83.3	806	1 ENPL_HORVU	P36183 hordeum vul
2	34	81.0	209	1 C551_CHLVI	P42426 chlorobium
3	33	78.6	118	1 R18E_SULTO	O96YV1 sulfolobus
4	33	78.6	659	1 VARI_PYRAB	O9UXU2 pyrococcus
5	32	76.2	325	1 REP9_ECOLI	O51651 escherichia
6	32	76.2	325	1 REP8_ECOLI	O57154 escherichia
7	32	76.2	325	1 RP10_ECOLI	O52219 escherichia
8	32	76.2	325	1 RP11_ECOLI	O52347 escherichia
9	32	76.2	325	1 RP12_SALTI	O57481 salmonella
10	31	73.8	312	1 CC2D_ANTMA	O38775 antirrhinum
11	31	73.8	412	1 SDAC_HAETN	P44615 haemophilus
12	31	73.8	794	1 PMS1_SCHPO	P54280 schizosacch
13	31	73.8	817	1 ENPL_CATRO	P35016 cathepsin
14	31	73.8	1709	1 CHD1_HUMAN	O14646 homo sapien
15	31	73.8	1711	1 CHD1_MOUSE	P40201 mus musculu
16	31	73.8	1739	1 CHD2_HUMAN	O14647 homo sapien
17	31	73.8	1944	1 CHD3_HUMAN	O12873 homo sapien
18	31	73.8	4473	1 PLE1_CRIGR	O9J135 cricetus
19	31	73.8	4684	1 PLE1_HUMAN	O15149 homo sapien
20	31	73.8	4687	1 PLE1_RAT	P30427 rattus norv
21	30	71.4	136	1 YAGP_ECOLI	P75684 escherichia
22	30	71.4	297	1 RGG_STRGC	P49330 streptococc
23	30	71.4	307	1 PANC_LOTJA	O24035 lotus japon
24	30	71.4	437	1 EFTU_YEAST	P02992 saccharomyc
25	30	71.4	591	1 PAK4_HUMAN	O96013 homo sapien
26	30	71.4	849	1 SRK6_BRAOL	O09092 brassica ol
27	30	71.4	925	1 NEC3_HUMAN	O13469 homo sapien
28	30	71.4	1442	1 PNC1_CHICK	O90693 gallus gall
29	29	69.0	186	1 Y450_METJA	O57892 methanococ
30	29	69.0	198	1 SARA_MOUSE	O9n131 homo sapien
31	29	69.0	198	1 SARA_MOUSE	P36536 mus musculu
32	29	69.0	217	1 CATL_SHEEP	O10991 ovis aries
33	29	69.0	218	1 CATL_CHICK	P09648 gallus gall

34	29	69.0	249	1 PYR2_FREDI	P18543 fremyella d
35	29	69.0	309	1 CC2B_HUMAN	P25859 arabidopsis
36	29	69.0	313	1 O2B6_HUMAN	P58173 homo sapien
37	29	69.0	334	1 CATL_BOVIN	P25975 bos taurus
38	29	69.0	334	1 CATL_PIG	O28944 sus scrofa
39	29	69.0	334	1 CS12_HUMAN	O60911 homo sapien
40	29	69.0	356	1 VAOD_DICDI	P54641 dictyosteli
41	29	69.0	397	1 PCK_STRPY	P82487 streptococc
42	29	69.0	449	1 Y753_SYNT3	P74635 synecocyst
43	29	69.0	451	1 NMT_CANAL	P30418 candida alb
44	29	69.0	474	1 ATPB_RHORU	P05038 rhodospiril
45	29	69.0	476	1 TH1L_THERVO	O97ak6 thermoplasma

## ALIGNMENTS

RESULT 1	ID	ENPL_HORVU	STANDARD:	PRT:	809 AA.
AC	P36183:	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)				
DT	01-FEB-1995 (Rel. 31, Last annotation update)				
DE	Endoplasmic homolog precursor (GRP94 homolog).				
OS	Hordeum vulgare (Barley).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;				
OC	Triflicae; Hordeum.				
OX	NCBI_TaxID=4513;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=CV. P-01; TISSUE=Leaf;				
RX	MEDLINE=93257625; PubMed=8490130;				
RA	Walther-Larsen H., Brandt J., Collinge D.B.,				
RA	Thorald-Christensen H.;				
RT	"A pathogen-induced gene of barley encodes a HSP90 homologue showing				
RT	striking similarity to vertebrate forms resident in the endoplasmic				
RT	reticulum."				
RL	Plant Mol. Biol. 21:1097-1108(1993).				
CC	- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.				
CC	- INDUCTION: ACCUMULATES RAPIDLY IN LEAVES UPON HEAT SHOCK				
CC	TREATMENT AND DURING INFECTION BY A PATHOGEN.				
CC	- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
CC	EMBL: X67960; CAA48143.1; -				
CC	PIR: S31862; S31862.				
CC	PIR: S33533; S33533.				
CC	HSP: P02829; IAH8.				
CC	InterPro: IPR003594; AtPbind_Atpase.				
CC	InterPro: IPR000886; ER_target.				
CC	InterPro: IPR001404; HSP90.				
CC	Pfam: PF00183; HSP90; 1.				
CC	Pfam: PF02518; HATPase_C; 1.				
CC	PRINTS: PR00775; HEATSHOCK90.				
CC	SMART: SM00387; HATPase_C; 1.				
CC	PROSITE: PS00014; ER_TARGET; 1.				
CC	PROSITE: PS00298; HSP90; 1.				
CC	Chaperone; Endoplasmic reticulum; Glycoprotein; Calcium-binding;				
CC	Signal.				
CC	FT SIGNAL	1	18		POTENTIAL.
CC	FT CHAIN	19	809		ENDOPLASMIN HOMOLOG.
CC	FT CAROHYD	111	111		N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	FT CAROHYD	410	410		N-LINKED (GLCNAC. . .) (POTENTIAL).
CC	FT CAROHYD	450	450		N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 806 809 PREVENT SECRETION FROM ER.  
 SQ SEQUENCE 809 AA; 92916 MW; 79798FDBEC15B44D0 CRC64;

Query Match 83.3%; Score 35; DB 1; Length 809;  
 Best Local Similarity 77.8%; Pred. No. 9;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
 DB 543 KSGDKLVSL 551

## RESULT 2

C551\_CHELY  
 ID C551\_CHELY STANDARD; PRT: 206 AA.  
 AC P42426;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Photosynthetic reaction center cytochrome C-551 (C551).  
 GN CYC OR PSCC.  
 OS Chlorobium vibrioforme.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorodiu.  
 ON NCBI\_TaxID=1098;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-33.  
 RC STRAIN=f. thiosulfatophilum / NCIB 8327;  
 RX MEDLINE=93016035; Pubmed=183218;  
 RA Okreks J.S., Kjaer B., Hansson O., Svendsen I., Moeller B.L.,  
 RA Scheller H.V.;  
 RT "A membrane-bound monoheme cytochrome c551 of a novel type is the  
 RT photosynthetic reaction center complex.";  
 RT J. Biol. Chem. 267:21139-21145(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=f. thiosulfatophilum / Larsen;  
 RA Oh-Oka H.;

RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MONOHEME CYTOCHROME WHICH IS THE IMMEDIATE ELECTRON  
 CC DONOR TO P840 OF THE PHOTOSYNTHETIC REACTION CENTER COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (probable).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M95751; AAA23110.1; -;  
 DR EMBL; AB004459; BAA20401.1; -;  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR PROSITE; PS00190; CYTOCHROME\_C\_1.  
 KW Electron transport; Photosynthesis; Reaction center; Heme;  
 KW Transmembrane.  
 FT TRANSMEM 10 30 POTENTIAL.  
 FT TRANSMEM 49 69 POTENTIAL.  
 FT TRANSMEM 76 96 POTENTIAL.  
 FT BINDING 152 152 HEME (COVALENT) (BY SIMILARITY).  
 FT BINDING 155 155 HEME (COVALENT) (BY SIMILARITY).  
 FT METAL 156 156 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 FT METAL 182 182 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 206 AA; 22858 MW; 56342FE9BA7AB82B CRC64;

Query Match 81.0%; Score 34; DB 1; Length 206;  
 Best Local Similarity 77.8%; Pred. No. 3.5;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10

Db 4 KSGKRLIAL 12

## RESULT 3

R18E\_SULTO  
 ID R18E\_SULTO STANDARD; PRT: 118 AA.  
 AC Q96YWL;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L18e.  
 GN RPL18E OR ST206.  
 OS Sulfolobus tokodaii.  
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
 OC Sulfolobus.  
 ON NCBI\_TaxID=11955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JCM 10545 / 7;  
 RX MEDLINE=21456156; Pubmed=11572479;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
 RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Oguchi A.,  
 RA Aoki K.-I., Masuda S., Yanagita M., Nishimura M., Yamagishi A.,  
 RA Oshima T., Kikuchi H.;

RT "Complete genome sequence of an aerobic thermacidophilic  
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
 RT DNA Res. 8:123-140(2001).  
 RL [1]  
 CC -1- SIMILARITY: BELONGS TO THE L18E FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AP000988; BAB67165.1; -;  
 DR InterPro; IPR001196; Ribosomal\_L15.  
 DR Pfam; PF00256; L15; 1.  
 DR PROSITE; PS01106; RIBOSOMAL\_L18E; FALSE\_NEG.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 118 AA; 13335 MW; CE18B01FE2F0B394 CRC64;

Query Match 78.6%; Score 33; DB 1; Length 118;  
 Best Local Similarity 77.8%; Pred. No. 3.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
 DB 92 KSGKRVISL 100

## RESULT 4

VATI\_PYRAB  
 ID VATI\_PYRAB STANDARD; PRT: 659 AA.  
 AC Q9UXU2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE V-type ATP synthase subunit I (EC 3.6.3.14) (V-type ATPase subunit I).  
 GN APTI OR PAB180.  
 OS Pyrococcus abyssi.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 ON NCBI\_TaxID=29292;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GES / Orsay;  
 RA Hellig R.;

```

RT      "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT      structure and evolution.";
RL      Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC      GRADIENT ACROSS THE MEMBRANE.
CC      -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC      H(+)(Out).
CC      -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AJ248288; CAB50671.1; -.
DR      InterPro; IPR002490; V_ATPase_sub116.
DR      Pfam; PF01496; V_ATPase_sub.a; 1.
KW      Hydrolase; Hydrogen ion transport; Transmembrane; Complete proteome.
SQ      SEQUENCE 659 AA; 74289 MW; DED5A21A8D58D740 CRC64;

Query Match
Best Local Similarity 77.8%; Score 33; DB 1; Length 659;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      2 KSSGKRLSL 10
       1 1111111
Db      206 KDSGKVIYL 214

RESULT 5
REP9_ECOLI STANDARD; PRT; 325 AA.
ID      REP9_ECOLI
AC      Q51651;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      RepFib replication protein A.
GN      RepA.
OS      Escherichia coli.
OC      Plasmid IncFI ColV.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=93361574; PubMed=8356112;
RA      Gibbs M.D., Spiers A.J., Bergquist P.L.;
RT      "RepFib: a basic replicon of large plasmids.";
RL      Plasmid 29.165-179(1993).
CC      -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT
CC      IS INVOLVED IN COPY CONTROL FUNCTIONS. IN VITRO, BINDS TO THE DNA
CC      REPEAT UNITS, BCDD'D'', EFG AND HIJ (BY SIMILARITY).
CC      -1- SIMILARITY: HIGH, TO REPFIB REPLICON REPA PROTEINS FROM OTHER
CC      PLASMIDS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L01250; AAA71880.1; -.
DR      InterPro; IPR002584; RepA.
DR      Pfam; PF01651; RepA; 1.
DR      ProDom; PD003734; RepA; 1.
KW      Plasmid; DNA replication; Plasmid copy control; DNA-binding.
SQ      SEQUENCE 325 AA; 37298 MW; B7B0B7761A9327DB CRC64;

```

```

Query Match
Best Local Similarity 66.7%; Score 32; DB 1; Length 325;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY      2 KSSGELVTL 11
       11111111
Db      3 KSSGELVTL 11

RESULT 6
REP9_ECOLI STANDARD; PRT; 325 AA.
ID      REP9_ECOLI
AC      Q57154; Q60231;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      RepFib replication protein A.
DE      REPA OR REPA.
OS      Escherichia coli.
OC      Plasmid F, Plasmid IncFI P307, Plasmid IncFI PH502, and
OC      Plasmid IncFI ColVbtrp.
OC      Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC      Escherichia.
OX      NCBI_TaxID=562;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=K12 / CR63; PLASMID=F;
RA      Shimizu H., Saitoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;
RT      "Complete nucleotide sequence of the F plasmid: its implications for
RT      organization and diversification of the plasmid genomes.";
RL      Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      PLASMID=IncFI PH502, and IncFI ColVbtrp;
RL      MEDLINE=93361574; PubMed=8356112;
RA      Gibbs M.D., Spiers A.J., Bergquist P.L.;
RT      "RepFib: a basic replicon of large plasmids.";
RL      Plasmid 29.165-179(1993).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      PLASMID=IncFI P307;
RL      MEDLINE=89213960; PubMed=2651415;
RA      Saul D., Spiers A.J., McNulty J., Gibbs M.G., Bergquist P.L.,
RA      Hill D.F.;
RT      "Nucleotide sequence and replication characteristics of RepFib, a
RT      basic replicon of IncF plasmids.";
RL      J. Bacteriol. 171:2697-2707(1989).
RN      [4]
RP      EXPRESSION REGULATION, AND DNA REPEAT ELEMENT BINDING.
RX      MEDLINE=93308080; PubMed=8320218;
RA      Spiers A.J., Bhana N., Bergquist P.L.;
RT      "Regulatory interactions between RepA, an essential replication
RT      protein, and the DNA repeats of RepFib from plasmid P307.";
RL      J. Bacteriol. 175:4016-4024(1993).
CC      -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT
CC      IS INVOLVED IN COPY CONTROL FUNCTIONS. IN VITRO, BINDS TO THE DNA
CC      REPEAT UNITS, BCDD'D'', EFG AND HIJ.
CC      -1- SIMILARITY: HIGH, TO REPFIB REPLICON REPA PROTEINS FROM OTHER
CC      PLASMIDS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AP001918; BA97903.1; -.
DR      EMBL; L01253; AAA71885.1; -.
DR      EMBL; L01256; AAA71886.1; -.
DR      EMBL; M26308; AAB61762.1; ALT_INIT.
KW      EcoGene; EC40046; repB.

```

DR InterPro: IPR002584; RepA.  
 DR Pfam: PF01651; RepA; 1.  
 DR ProDom: PD003734; RepA; 1.  
 KW Plasmid; DNA replication; Plasmid copy control; DNA-binding;  
 KW Complete proteome.  
 FT VARIANT 60 60 A -> D (IN COPY NUMBER MUTANT).  
 SQ SEQUENCE 325 AA; 37283 MW; B7A4B637B18997DB CRC64;

Query Match 76.2%; Score 32; DB 1; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLSL 10  
 DB 3 KSSGELVLT 11

## RESULT 7

ID RP10\_ECOLI STANDARD; PRT; 325 AA.  
 AC 052219;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE RepFIB replication protein A.  
 GN RepA.  
 OS Escherichia coli.  
 OG Plasmid IncFI PH8507.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93361574; PubMed=8356112;  
 RA Gibbs M.D., Spiers A.J., Bergquist P.L.;  
 RT "RepFIB: a basic replicon of large plasmids.";  
 RL Plasmid 29:165-179(1993).  
 CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT  
 IS INVOLVED IN COPY CONTROL FUNCTIONS. IN VITRO, BINDS TO THE DNA  
 REPEAT UNITS, BCD'D'', EFG AND HIJ (BY SIMILARITY).  
 CC -!- SIMILARITY: HIGH, TO REPFIB REPLICON REPA PROTEINS FROM OTHER  
 PLASMIDS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 CC EMBL, L01252; AAA71882.1; -  
 DR InterPro: IPR002584; RepA.  
 DR Pfam: PF01651; RepA; 1.  
 DR ProDom: PD003734; RepA; 1.  
 KW Plasmid; DNA replication; Plasmid copy control; DNA-binding.  
 SQ SEQUENCE 325 AA; 37249 MW; B7A4BC3DB1839DDB CRC64;

Query Match 76.2%; Score 32; DB 1; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLSL 10  
 DB 3 KSSGELVLT 11

## RESULT 8

ID RP11\_ECOLI STANDARD; PRT; 325 AA.  
 AC 052347;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE RepFIB replication protein A.  
 GN RepA.  
 OS Escherichia coli.  
 OG Plasmid IncFI R386.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.  
 NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93361574; PubMed=8356112;  
 RA Gibbs M.D., Spiers A.J., Bergquist P.L.;  
 RT "RepFIB: a basic replicon of large plasmids.";  
 RL Plasmid 29:165-179(1993).  
 CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT  
 IS INVOLVED IN COPY CONTROL FUNCTIONS. IN VITRO, BINDS TO THE DNA  
 REPEAT UNITS, BCD'D'', EFG AND HIJ (BY SIMILARITY).  
 CC -!- SIMILARITY: HIGH, TO REPFIB REPLICON REPA PROTEINS FROM OTHER  
 PLASMIDS.

CC -----  
 CC EMBL, L01254; AAA71884.1; -  
 DR InterPro: IPR002584; RepA.  
 DR Pfam: PF01651; RepA; 1.  
 DR ProDom: PD003734; RepA; 1.  
 KW Plasmid; DNA replication; Plasmid copy control; DNA-binding.  
 SQ SEQUENCE 325 AA; 37214 MW; 3C1EC7EC5B8C0870 CRC64;

Query Match 76.2%; Score 32; DB 1; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLSL 10  
 DB 3 KSSGELVLT 11

## RESULT 9

ID RP12\_SALTI STANDARD; PRT; 325 AA.  
 AC 057481;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE RepFIB replication protein A.  
 GN RepA.  
 OS Salmonella typhi, and  
 OS Escherichia coli.  
 OG Plasmid IncFIV R124, and plasmid IncFI CO1V3-K30.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmonella.

CC -----  
 CC NCB1\_TaxID=601, 562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93361574; PubMed=8356112;  
 RA Gibbs M.D., Spiers A.J., Bergquist P.L.;  
 RT "RepFIB: a basic replicon of large plasmids.";  
 RL Plasmid 29:165-179(1993).  
 CC -!- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR PLASMID REPLICATION; IT  
 IS INVOLVED IN COPY CONTROL FUNCTIONS. IN VITRO, BINDS TO THE DNA  
 REPEAT UNITS, BCD'D'', EFG AND HIJ (BY SIMILARITY).  
 CC -!- SIMILARITY: HIGH, TO REPFIB REPLICON REPA PROTEINS FROM OTHER  
 PLASMIDS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -



CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L01251; AAA71881.1; -  
 DR EMBL; L01251; AAA71881.1; -  
 DR InterPro: IPR002584; RepA.  
 DR Pfam: PF01651; RepA; 1.  
 DR ProDom: PD003734; RepA; 1.  
 KW Plasmid; DNA replication; Plasmid copy control; DNA-binding.  
 SQ SEQUENCE 325 AA; 37315 MW; B7A4B6361A9327DB CRC64;

Query Match 76.2%; Score 32; DB 1; Length 325;  
 Best Local Similarity 66.7%; Pred. No. 16;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
 |||||:  
 DB 3 KSSGELVTL 11

RESULT 10  
 ID CC2D\_ANTMA STANDARD; PRT; 312 AA.  
 AC 038775;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Cell division control protein 2 homolog D (EC 2.7.1.-).  
 GN CC2D.  
 OS *Escherichia majus* (Garden snapperagon).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.  
 OX NCBI\_Taxid=4151;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=flower;  
 RX MEDLINE=96434535; PubMed=8837502;  
 RA Fobert P.R., Gaudin V., Lunness P., Coen E.S., Doonan J.H.;  
 RT "distinct classes of cdc2-related genes are differentially expressed  
 RT during the cell division cycle in plants.";  
 RL Plant Cell 8:1465-1476(1996).  
 CC -1- FUNCTION: PLAYS A KEY ROLE IN THE CONTROL OF THE EUKARYOTIC  
 CC CELL CYCLE.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION SPECIFIC TO THE G2 AND M PHASES.  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC CDC2/CDK SUBFAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X97640; CA66236.1; -  
 DR HSHP; P24941; 1B38.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR002290; Ser\_thr\_pkinase.  
 DR Pfam: PF00069; Pkinase; 1.  
 DR ProDom: PD000001; Euk\_Pkinase; 1.  
 DR SMART; SM00220; S\_TKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Transferrase; Serine/threonine-protein kinase; ATP-binding;  
 KW Cell cycle; Cell division; Mitosis; Phosphorylation.  
 FT DOMAIN 14 304 PROTEIN KINASE.  
 FT NP\_BIND 20 28 ATP (BY SIMILARITY).

FT BINDING 43 43 ATP (BY SIMILARITY).  
 FT ACT\_SITE 145 145 BY SIMILARITY.  
 FT MOD\_RES 24 24 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 25 25 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 179 179 PHOSPHORYLATION (BY CAK) (BY SIMILARITY).  
 SQ SEQUENCE 312 AA; 35503 MW; A38D2177BEF70210 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 312;  
 Best Local Similarity 55.6%; Pred. No. 25;  
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 2 KSSGKLISL 10  
 |||||:  
 DB 34 KSTGRVAL 42

RESULT 11  
 ID SDAC\_HAEIN STANDARD; PRT; 412 AA.  
 AC P44615;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serine transporter.  
 GN SDAC OR HI0289.  
 OS *Haemophilus influenzae*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_Taxid=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Rd / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: INVOLVED IN THE IMPORT OF SERINE INTO THE CELL  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE SDAC/TDC FAMILY OF TRANSPORTERS.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; U32715; AAC21954.1; -  
 DR TIGR; HI0289; -  
 DR InterPro: IPR002422; AA/rel\_pmease2.  
 DR InterPro: IPR004694; Ser\_transp.  
 DR TIGRFAMs: TIGR00814; stp; 1.  
 KW Transport; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 31 51 POTENTIAL.  
 FT TRANSMEM 91 111 POTENTIAL.  
 FT TRANSMEM 131 151 POTENTIAL.  
 FT TRANSMEM 155 175 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 240 260 POTENTIAL.  
 FT TRANSMEM 286 306 POTENTIAL.  
 FT TRANSMEM 338 358 POTENTIAL.

```

FT TRANSMEM 361 381 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
SO SEQUENCE 412 AA; 45894 MW; 147CIDAZP090008 CRC64;

Query Match 73.8%; Score 31; DB 1; Length 412;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10
Db 86 KTACKLITL 94

RESULT 12
PMSL_SCHPO STANDARD; PRT; 794 AA.
AC P54280;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA mismatch repair protein pms1.
PM1 OR SPAC19G12.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=97403304; PubMed=9258673;
RA Schar P., Baur M., Schneider C., Kohli J.;
RT "Mismatch repair in Schizosaccharomyces pombe requires the mult.
RL homologous gene pms1: molecular cloning and functional analysis.";
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkert G., Aert R., Robben J., Grymonprez B.,
RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moesti D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaur V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Carrilli L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES
CC IN DNA.
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEX FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X96581; CA65400.1; -
DR EMBL; Z97209; CAB1013.1; -
DR HSSP; P23367; IBKN.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR002099; DNA_mis_repair.
DR Pfam; PF01119; DNA_mis_repair.
DR Pfam; PF02518; HATPase_c; 1.
DR TIGRFAMS; TIGR00585; mult; 1.
DR PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
KW DNA repair.
SO SEQUENCE 794 AA; 88009 MW; A5D46FFFA077D8DC CRC64;

Query Match 73.8%; Score 31; DB 1; Length 794;
Best Local Similarity 66.7%; Pred. No. 68;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10
Db 197 KNSGKLQQL 205

RESULT 13
ENPL_CATRO STANDARD; PRT; 817 AA.
ID ENPL_CATRO
AC P35016;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endoplasmic reticulum precursor (GRP94 homolog).
GN HSP90.
OS Catharanthus roseus (rosy periwinkle) (Madagascar periwinkle).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Gentianales; Apocynaceae; Rauvolfioideae;
OC Vinceae; Catharanthus.
OX NCBI_TaxID=4058;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. GP3A;
RX MEDLINE=9403337; PubMed=8106014;
RA Schroeder G., Beck M., Eichel J., Vetter H.P., Schroeder J.;
RT "HSP90 homologue from Madagascar periwinkle (Catharanthus roseus):
RT endoplasmic reticulum.";
RT Plant Mol. Biol. 23:583-594(1993).
CC -1- FUNCTION: MAY HAVE A MOLECULAR CHAPERONE ROLE IN THE PROCESSING OF
CC SECRETED MATERIALS. IT IS PROBABLY NOT A MAJOR HEAT SHOCK RESPONSE
CC COMPONENT.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- TISSUE SPECIFICITY: NOT DETECTED IN EXTRACTS FROM YOUNG PLANTS
CC UNLESS THEY ARE EXPOSED TO HEAT SHOCK FOR SEVERAL HOURS.
CC FOUND TO BE CONSTITUTIVELY EXPRESSED IN CELL CULTURES.
CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L14594; AAA16785.1; -
DR PIR; S39558; S39558.
DR HSSP; P02829; 1AH8.
DR InterPro: IPR003594; ATPbind_ATPase.
DR InterPro: IPR000886; ER_target.
DR InterPro: IPR001404; Hsp90.
DR Pfam; PF00183; HSP90; 1.

```

DR Pfam: PF02518; HATPase\_C.1.  
 DR PRINTS: PR00775; HEATSHOCK90.  
 DR SMART: SM00387; HATPase\_C.1.  
 DR PROSITE: PS00014; ER\_TARGET.1.  
 DR PROSITE: PS00298; HSP90.1.  
 DR Chapterone: Endoplasmic reticulum; Glycoprotein; Calcium-binding;  
 signal.  
 KM SIGNAL.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 817 ENDOPLASMIC HOMOLOG.  
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 814 817 PREVENT SECRETION FROM ER.  
 SO SEQUENCE 817 AA; 93491 MW; 26C06CDC5EDD19FA CRC64;  
 Query Match 73.8%; Score 31; DB 1; Length 817;  
 Best Local Similarity 77.8%; Pred. No. 70;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 Oy 2 KSSGKLISL 10  
 Db 550 KSEGLTSL 558

RESULT 14  
 CHD1\_HUMAN STANDARD; PRT; 1709 AA.  
 AC 014646;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chromodomain-helicase-DNA-binding protein 1 (CHD-1).  
 GN CHD1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=97470991; PubMed=9326634;  
 RA Woodage T., Basrai M.A., Baxevanis A.D., Hieter P., Collins F.S.;  
 RT "Characterization of the CHD family of proteins."  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11472-11477(1997).  
 CC -1- FUNCTION: SEQUENCE-SELECTIVE DNA-BINDING PROTEIN. COULD PLAY AN  
 IMPORTANT ROLE IN GENE REGULATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 CHROMO DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF006513; AB87381.1; -  
 DR HSSP: P23197; IAP0.  
 DR Genew: HGNC:1915; CHD1.  
 DR MIM: 602118; -  
 DR InterPro: IPR000953; Chromo.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR000330; SNF2\_N.  
 DR Pfam: PF00176; SNF2\_N.1.  
 DR Pfam: PF00271; helicase\_C.1.  
 DR Pfam: PF00385; chromo; 2.  
 DR SMART: SM00298; CHROMO; 2.  
 DR SMART: SM00487; DEXDC; 1.  
 DR SMART: SM00490; HELICG; 1.

DR PROSITE: PS00598; CHROMO\_1; 2.  
 DR PROSITE: PS50013; CHROMO\_2; 2.  
 KM DNA-binding; ATP-binding; Helicase; Nuclear protein; Repeat.  
 FT DOMAIN 1 70 SER-RICH.  
 FT DOMAIN 117 137 SER-RICH.  
 FT DOMAIN 272 364 CHROMO\_1.  
 FT DOMAIN 389 452 CHROMO\_2.  
 FT NP\_BIND 506 513 ATP (POTENTIAL).  
 FT SITE 614 617 DEAD BOX.  
 FT DOMAIN 1628 1644 3 X 5 AA REPEATS OF H-S-D-H-R.  
 FT REPEAT 1628 1632 1.  
 FT REPEAT 1634 1638 2.  
 FT REPEAT 1640 1644 3.  
 SO SEQUENCE 1709 AA; 196517 MW; 416409C913D6A935 CRC64;  
 Query Match 73.8%; Score 31; DB 1; Length 1709;  
 Best Local Similarity 77.8%; Pred. No. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 KSSGKLISL 10  
 Db 785 RSSGKLILL 793

RESULT 15  
 CHD1\_MOUSE STANDARD; PRT; 1711 AA.  
 AC P40201;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Chromodomain-helicase-DNA-binding protein 1 (CHD-1).  
 GN CHD1 OR CHD-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP MEDLINE=93211972; PubMed=8460153;  
 RA Delmas V., Stokes D.G., Perry R.P.;  
 RT "A mammalian DNA-binding protein that contains a chromodomain and an  
 SNF2/SWI2-like helicase domain."  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2414-2418(1993).  
 CC -1- FUNCTION: SEQUENCE-SELECTIVE DNA-BINDING PROTEIN. COULD PLAY AN  
 IMPORTANT ROLE IN GENE REGULATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: ABUNDANCE IS HIGHER IN CELLS REPRESENTING  
 EARLY STAGES OF THE B LYMPHOID LINEAGE SUCH AS PRE-B AND B CELLS,  
 THAN IN CELLS REPRESENTING MATURE PLASMACYTES OR OTHER CELL  
 LINESAGES SUCH AS FIBROBLASTS.  
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 CHROMO DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L10410; AA808486.1; -  
 DR PIR: A47392; A47392.  
 DR HSSP: P23197; IAP0.  
 DR MGD: MGI:88393; Chd1.  
 DR InterPro: IPR000953; Chromo.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR000330; SNF2\_N.  
 DR Pfam: PF00176; SNF2\_N.1.  
 DR Pfam: PF00271; helicase\_C.1.  
 DR Pfam: PF00385; chromo; 2.

DR SMART; SM00298; CHROMO; 2.  
 DR SMART; SM00487; DEXDC; 1.  
 DR SMART; SM00490; HELIC; 1.  
 DR PROSITE; PS00598; CHROMO\_1; 2.  
 DR PROSITE; PS00013; CHROMO\_2; 2.  
 KM DNA-binding; ATP-binding; Helicase; Nuclear protein; Repeat.  
 FT DOMAIN 1 70 SER-RICH.  
 FT DOMAIN 116 136 SER-RICH.  
 FT DOMAIN 270 362 CHROMO 1.  
 FT DOMAIN 387 450 CHROMO 2.  
 FT NP\_BIND 504 511 ATP (POTENTIAL).  
 FT SITE 612 615 DEAD BOX.  
 FT DOMAIN 1629 1645 3 X 5 AA REPEATS OF H-S-D-H-R.  
 FT REPEAT 1629 1633 1.  
 FT REPEAT 1635 1639 2.  
 FT REPEAT 1641 1645 3.  
 SQ SEQUENCE 1711 AA; 196409 MW; FE3F8FD13E32E24 CRC64;

Query Match 73.88; Score 31; DB 1; Length 1711;  
 Best Local Similarity 77.88; Pred. NO. 1.5e+02;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSGKLSL 10  
 :|||||  
 Db 783 RSSKLLIL 791

Search completed: January 10, 2003, 09:38:40  
 Job time : 12 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:15 ; Search time 29 seconds  
(without alignments)  
78.156 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSSGKLISLX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp.\_archaea:\*  
2: sp.\_bacteria:\*  
3: sp.\_fungi:\*  
4: sp.\_human:\*  
5: sp.\_invertebrate:\*  
6: sp.\_mammal:\*  
7: sp.\_mhc:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_rodent:\*  
12: sp.\_virus:\*  
13: sp.\_vertebrate:\*  
14: sp.\_unclassified:\*  
15: sp.\_vivirus:\*  
16: sp.\_bacteriaph:\*  
17: sp.\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

No.	Score	Query	Length	ID	Description
1	37	88.1	763	10	Q9FZ11
2	34	81.0	88	16	Q97MX6
3	34	81.0	1465	5	O17909
4	33	78.6	192	5	Q9V43
5	33	78.6	559	10	Q9M015
6	33	78.6	650	13	Q90M62
7	32	76.2	116	2	P97240
8	32	76.2	127	2	P71249
9	32	76.2	127	2	P71250
10	32	76.2	137	2	P75032
11	32	76.2	127	2	P75032
12	32	76.2	247	16	Q97MH2
13	32	76.2	256	11	Q9PUC5
14	32	76.2	298	11	Q9D3J9
15	32	76.2	301	4	Q9HA16
16	32	76.2	324	5	Q9Y0D2

17	32	76.2	325	2	Q93QL0	Q93QL0 escherichia
18	32	76.2	325	2	Q8VU37	Q8VU37 escherichia
19	32	76.2	343	10	Q9FS02	Q9FS02 astragalus
20	32	76.2	389	10	Q9L0T9	Q9L0T9 arabidopsis
21	32	76.2	706	1	Q9C4L4	Q9C4L4 thermococcus
22	32	76.2	766	10	Q9FJ62	Q9FJ62 arabidopsis
23	32	76.2	856	10	Q39277	Q39277 brassica ca
24	31	73.8	218	16	Q9X279	Q9X279 thermotoga
25	31	73.8	264	17	Q8TPQ0	Q8TPQ0 methanosaeta
26	31	73.8	292	2	Q45422	Q45422 bacillus sp
27	31	73.8	305	5	Q19986	Q19986 caenorhabditis
28	31	73.8	312	11	Q8VE72	Q8VE72 mus musculus
29	31	73.8	313	11	Q8VC87	Q8VC87 mus musculus
30	31	73.8	317	10	Q9M555	Q9M555 euphorbia e
31	31	73.8	320	16	Q92772	Q92772 chlamydia p
32	31	73.8	335	10	Q22972	Q22972 arabidopsis
33	31	73.8	349	15	Q71129	Q71129 human immun
34	31	73.8	350	15	Q71130	Q71130 human immun
35	31	73.8	425	12	Q8UY68	Q8UY68 simian aden
36	31	73.8	473	11	Q9D5K6	Q9D5K6 mus musculus
37	31	73.8	560	13	Q08781	Q08781 gallus gall
38	31	73.8	565	10	Q9FEE6	Q9FEE6 oryza sativ
39	31	73.8	631	16	Q8Z0U8	Q8Z0U8 salmonella
40	31	73.8	672	16	P73021	P73021 synecocyst
41	31	73.8	743	4	Q9HCJ5	Q9HCJ5 homo sapien
42	31	73.8	797	10	Q9CAX2	Q9CAX2 arabidopsis
43	31	73.8	810	10	Q9MB32	Q9MB32 oryza sativ
44	31	73.8	812	10	Q8SB39	Q8SB39 oryza sativ
45	31	73.8	823	10	Q9STX5	Q9STX5 arabidopsis

## ALIGNMENTS

RESULT 1	
Q9FZ11	PRELIMINARY; PRT; 763 AA.
ID Q9FZ11	
AC Q9FZ11	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE F1019.5 protein.	
OS F1019.5.	
GN Arabidopsis thaliana (Mouse-ear cress).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.	
OX NCBI_Taxid=3702;	
RN [1]	
RP SEQUENCE FROM N.A.	
RA Federici N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,	
RA Altail H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,	
RA Buehler E., Chao O., Chiu J., Choi E., Gonzalez A.,	
RA Huang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,	
RA Lenz C., Liu A., Liu S., Mukharly N., Pham P., Sakano H., Shinn P.,	
RA Toriumi M., Vayberg M., Yu G., Ecker J., Theologis A., Davis R.W.,	
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AC007152; AAF98209.1; -	
DR InterPro; IPR004129; GDPD.	
DR Pfam; PFO3009; GDPD; 1.	
SQ SEQUENCE 763 AA; 83787 MW; CFFAD6E6334D30BB CRC64;	
Query Match	88.1%; Score 37; DB 10; Length 763;
Best Local Similarity	88.9%; Pred. No. 21;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY 2 KSSGKLISL 10	
Db 477 KNSGKLISL 485	

RESULT 2

Q97MX6

```

ID 097MX6 PRELIMINARY; PRT; 88 AA.
AC 097MX6;
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Hypothetical protein CAC0064.
GN CAC0064.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Bacilli; Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Oiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabaite F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL: AE007519; AAK78050.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10078 MW; 73D3D510DD44FE80 CRC64;

Query Match 81.0%; Score 34; DB 16; Length 88;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10
Db 17 KSSGEIISL 25

RESULT 3
017909 PRELIMINARY; PRT; 1465 AA.
ID 017909;
AC 017909;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE H6001.2 protein.
GN H6001.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Barlow K.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 "CHROMO" DOMAINS.
CC EMBL: Z92970; CAB07481.1; -.
DR InterPro: IPR000953; Chromo.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002464; DEAD box.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR000330; SNF2_N.
DR Pfam: PF00385; Chromo; 2.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00176; SNF2_N; 1.
DR SMART: SM00398; CHROMO; 2.
DR SMART: SM00487; DEADC; 1.
DR SMART: SM00490; HELICG; 1.

```

```

DR PROSITE; PS00598; CHROMO_1; 2.
DR PROSITE; PS50013; CHROMO_2; 2.
DR PROSITE; PS00690; DEAD ATP HELICASE; UNKNOWN_1.
KW ATP-binding; Helicase; Nuclear protein.
SQ SEQUENCE 1465 AA; 169391 MW; 87BE6FB5BDDC81 CRC64;

Query Match 81.0%; Score 34; DB 5; Length 1465;
Best Local Similarity 88.9%; Pred. No. 1,9e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10
Db 708 KSSGKLISL 716

RESULT 4
09VV43 PRELIMINARY; PRT; 192 AA.
ID 09VV43;
AC 09VV43;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CG4893 protein (RE55630P) (RE39465P).
GN CG4893.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sultón G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouch J., Brokstein P., Brotler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegian C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleib J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shier B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.W., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN-BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Abhayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phoumenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Seimster S.;  
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AE003527; AAF49479.1; -  
DR EMBL: AY071493; AAL49415.1; -  
DR EMBL: AY071358; AAL48960.1; -  
DR FLYBASE: FBgn0036616; CG4893.  
SQ SEQUENCE 192 AA; 20596 MW; 95654DBA81A3E50 CRC64;

Query Match 78.6%; Score 33; DB 5; Length 192;  
Best Local Similarity 77.8%; Pred. No. 40;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
Db 51 KSDGKLITL 59

RESULT 5  
ID Q9M015 PRELIMINARY; PRT; 559 AA.  
AC Q9M015;

DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Putative NADH dehydrogenase.  
GN ATG28220.

OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,  
RA Mewes H.W., Lemcke K., Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
CC -i COFACTOR: FAD (BY SIMILARITY).  
DR EMBL: AL161572; CAB79624.1; -

DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR001327; FAD\_pyr\_redox.  
DR InterPro: IPR00103; Pyridine\_redox\_2.  
DR InterPro: IPR001100; Pyr\_redox.  
DR Pfam: PF00036; efhand; 1.

DR Pfam: PF00070; Pyr\_redox; 1.  
DR PRINTS: PR00368; FADPNR.  
DR PRINTS: PR00411; PNDRDTASE1.  
DR PRINTS: PR00469; PNDRDTASE1.  
DR PROSITE: PS00018; EF\_HAND; UNKNOWN\_1.  
KW FAD; Flavoprotein; Oxidoreductase.

SQ SEQUENCE 559 AA; 61784 MW; ADC2DDED20EF18BA CRC64;  
Query Match 78.6%; Score 33; DB 10; Length 559;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
Db 305 KSSGELVSI 313

RESULT 6  
Q90W62 PRELIMINARY; PRT; 650 AA.  
ID Q90W62

AC Q90W62;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE PAK5 protein.  
GN PAK5.

OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;

RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-STAGE VI OOCYTE;  
RA Gau J., Faure S., Delser C., Morin N.;  
RT "A novel xenopus p21 activated kinase expressed in brain."  
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.

DR EMBL: AJ277826; CAC40979.1; -  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR00095; PAKbox/Rhodning.  
DR Pfam: PF00786; PBD; 1.

DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 1.  
DR PROSITE: PS50108; GBD; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.

KW ATP-binding; Transferase.  
SQ SEQUENCE 650 AA; 73736 MW; 9274DC6CACDA081 CRC64;  
Query Match 78.6%; Score 33; DB 13; Length 650;  
Best Local Similarity 66.7%; Pred. No. 1.4e+02;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
Db 400 KSSGKLIVAV 408

RESULT 7  
ID P97240 PRELIMINARY; PRT; 116 AA.  
AC P97240;

DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Strain SARA 63 serotype MOENCHEN F-related plasmid REPA (Fragment).  
GN REPA.

OS Salmonella typhimurium.  
OC Plasmid F-related.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OX NCBI\_TaxID=602;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SARA 63;  
RX MEDLINE=97197526; PubMed=9045822;  
RA Boyd E.F., Hartl D.L.;

RT "Recent horizontal transmission of plasmids between natural  
RL populations of Escherichia coli and Salmonella enterica".  
J. Bacteriol. 179:1622-1627(1997).  
DR EMBL: U81637; ABA49356.1; -

DR EMBL: U81635; ABA49354.1; -  
DR EMBL: U81636; ABA49355.1; -  
DR InterPro: IPR002584; REPA.  
DR Pfam: PF01651; REPA; 1.

DR ProDom: PD003734; REPA; 1.  
KW Plasmid.  
FT NON\_TER 1 1  
FT NON\_TER 116 116  
SQ SEQUENCE 116 AA; 12812 MW; F9183FD2B00EAA95 CRC64;  
Query Match 76.2%; Score 32; DB 2; Length 116;  
Best Local Similarity 66.7%; Pred. No. 39;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLSL 10  
||||:|:|:|  
Db 3 KSSGELVTL 11

## RESULT 8

P71249 PRELIMINARY; PRT; 127 AA.  
AC P71249;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPA (Fragment).  
GN REPA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ECOR11;  
RX MEDLINE=96400908; PubMed=8807284;  
RA Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;  
RT "Mosaic structure of plasmids from natural populations of Escherichia coli".  
RL Genetics 143:1091-1100(1996).  
DR EMBL: U50667; AAC44251.1; -;  
DR InterPro: IPR002584; REPA.  
DR Pfam: PF01651; REPA; 1.  
DR ProDom: PD003734; REPA; 1.  
FT NON\_TER 127  
SQ SEQUENCE 127 AA; 14310 MW; 90BB2994D52E37D0 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 127;  
Best Local Similarity 66.7%; Pred. No. 43;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLSL 10  
||||:|:|:|  
Db 3 KSSGELVTL 11

## RESULT 9

P71250 PRELIMINARY; PRT; 127 AA.  
AC P71250;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPA (Fragment).  
GN REPA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ECOR70;  
RX MEDLINE=96400908; PubMed=8807284;  
RA Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;  
RT "Mosaic structure of plasmids from natural populations of Escherichia coli".  
RL Genetics 143:1091-1100(1996).  
DR EMBL: U50677; AAC44261.1; -;  
DR InterPro: IPR002584; REPA.  
DR Pfam: PF01651; REPA; 1.  
DR ProDom: PD003734; REPA; 1.  
FT NON\_TER 127  
SQ SEQUENCE 127 AA; 14282 MW; 90BB298A35C027D0 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 127;

Best Local Similarity 66.7%; Pred. No. 43;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLSL 10  
||||:|:|:|  
Db 3 KSSGELVTL 11

## RESULT 10

P77639 PRELIMINARY; PRT; 127 AA.  
AC P77639;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPA (Fragment).  
GN REPA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ECOR48;  
RX MEDLINE=96400908; PubMed=8807284;  
RA Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;  
RT "Mosaic structure of plasmids from natural populations of Escherichia coli".  
RL Genetics 143:1091-1100(1996).  
DR EMBL: U50674; AAC44258.1; -;  
DR EMBL: U50673; AAC44257.1; -;  
DR InterPro: IPR002584; REPA.  
DR Pfam: PF01651; REPA; 1.  
DR ProDom: PD003734; REPA; 1.  
FT NON\_TER 127  
SQ SEQUENCE 127 AA; 14241 MW; 90BB28D1D52E37D0 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 127;  
Best Local Similarity 66.7%; Pred. No. 43;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLSL 10  
||||:|:~:~  
Db 3 KSSGELVTL 11

## RESULT 11

P75032 PRELIMINARY; PRT; 127 AA.  
AC P75032;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE REPA (Fragment).  
GN REPA.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VARIOUS STRAINS;  
RX MEDLINE=96400908; PubMed=8807284;  
RA Boyd E.F., Hill C.W., Rich S.M., Hartl D.L.;  
RT "Mosaic structure of plasmids from natural populations of Escherichia coli".  
RL Genetics 143:1091-1100(1996).  
DR EMBL: U50678; AAC44262.1; -;  
DR EMBL: U50668; AAC44252.1; -;  
DR EMBL: U50669; AAC44253.1; -;  
DR EMBL: U50670; AAC44254.1; -;  
DR EMBL: U50671; AAC44255.1; -;  
DR EMBL: U50672; AAC44256.1; -;



DR EMBL: U50675; AAC44259.1; -  
 DR EMBL: U50676; AAC44260.1; -  
 DR InterPro: IPR002584; Repa.  
 DR Pfam: PF01651; Repa. 1.  
 DR ProDom: PD003734; Repa. 1.  
 FT NON\_TER 127 127  
 SQ SEQUENCE 127 AA; 14213 MW; 90BB28CF35C027D0 CRC64;

Query Match 76.2%; Score 32; DB 2; Length 127;  
 Best Local Similarity 66.7%; Pred. No. 43;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10  
 |||||:  
 Db 3 KSSGELVTL 11

RESULT 12  
 097TH2 PRELIMINARY; PRT; 247 AA.

AC 097TH2;  
 DT 01-OCT-2001 (TREMblrel. 18, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Glycogen-binding regulatory subunit of S/T protein phosphatase  
 GN I.  
 OS Clostridium acetobutylicum.  
 OG Plasmid pS01.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 CC Clostridiales; Clostridiaceae; Clostridium.  
 OX NCBI\_TaxID=1488;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
 RX MEDLINE=21359325; PubMed=11466286;  
 RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
 RA Tatusov R.L., Sabahe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
 RA Bennett G.N., Koonin E.V., Smith D.R.;  
 RT "Genome sequence and comparative analysis of the solvent-producing  
 RT bacterium Clostridium acetobutylicum."  
 RL J. Bacteriol. 183:4823-4838(2001).  
 DR EMBL: AE001438; AKR76874.1; -  
 DR InterPro: IPR005036; CBM\_21.  
 DR Pfam: PF03370; CBM\_21; 2.  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 247 AA; 28026 MW; 68253021ED74BA27 CRC64;

Query Match 76.2%; Score 32; DB 16; Length 247;  
 Best Local Similarity 66.7%; Pred. No. 85;  
 Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10  
 |||||:  
 Db 5 KSAGKITFL 13

RESULT 13  
 09PUC5 PRELIMINARY; PRT; 256 AA.

AC 09PUC5;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Proteasome activator subunit 3.  
 GN PSME3.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20308156; PubMed=10849367;  
 RA Murray B.M., Sultmann H., Klein J.;  
 RT "Identification and linkage of the proteasome activator complex PA28  
 RT subunit genes in zebrafish."  
 RL Scand. J. Immunol. 51:571-576(2000).  
 DR EMBL: AF195050; AAF05816.1; -  
 DR HSSP: 006323; IAVO  
 DR ZFIN: ZDB-GENE-991110-19; psme3.  
 DR InterPro: IPR003185; PA28\_alpha.  
 DR InterPro: IPR003186; PA28\_beta.  
 DR Pfam: PF02251; PA28\_alpha. 1.  
 DR Pfam: PF02252; PA28\_beta. 1.  
 SQ SEQUENCE 256 AA; 29296 MW; 4FD4857849A35A55 CRC64;

Query Match 76.2%; Score 32; DB 13; Length 256;  
 Best Local Similarity 66.7%; Pred. No. 88;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10  
 |||||:  
 Db 112 KSNGLVEL 120

RESULT 14  
 09D3J9 PRELIMINARY; PRT; 298 AA.

AC 09D3J9;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE 5430427019Rik protein.  
 GN 5430427019Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=HEAD;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Kono H., Adachi Y., Fukuda S.,  
 RA Saito K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilting L.,  
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK017351; BAB30703.1; -  
 DR MGD: MGI:1918648; 5430427019Rik.  
 DR InterPro: IPR002198; ADH\_short.  
 DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.  
 SQ SEQUENCE 298 AA; 33057 MW; 4D4E47B146022BFB CRC64;

Query Match 76.2%; Score 32; DB 11; Length 298;  
 Best Local Similarity 87.5%; Pred. No. 1e+02; 1; Indels 0; Gaps 0;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KSSGKLIS 9  
 |||||:

Db 59 KSSGKFTS 66

## RESULT 15

09HA16 PRELIMINARY; PRT; 301 AA.  
 AC 09HA16; 01-MAR-2001 (TReMBLrel. 16, Created)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Hypothetical 33.9 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=EMBRO;  
 RA Iisogal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Magatsuma M., Hosoi T., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,  
 RA Niinomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=PLACENTA;  
 RA Strausberg R.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK021639; BAB13860.1; -;  
 DR EMBL; BC020611; AAH20611.1; -;  
 DR InterPro: IPR002198: ADH\_short.  
 DR PROSITE: PS00061: ADH\_SHORT; UNKNOWN\_1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 301 AA; 33894 MW; 0754B491077927C4 CRC64;

## Query Match

Best Local Similarity 76.28; Score 32; DB 4; Length 301;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKFTS 9

|||||  
 Db 59 KSSGKFTS 66

Search completed: January 10, 2003, 09:39:17  
 Job time : 32 secs

Fri, Jan 10 12:21:24 2003

us-09-869-003-1.rapm

Stuckler  
09/869003  
Seq. ID 1 by  
Inters Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:38:31 : Search time 379 Seconds  
(without alignments)  
18.713 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 XKSSGKLISLX 11

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/1/paa/PCrUS.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US081.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US082.COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US083.COMB.pep:\*  
8: /cgn2\_6/ptodata/1/paa/US084.COMB.pep:\*  
9: /cgn2\_6/ptodata/1/paa/US085.COMB.pep:\*  
10: /cgn2\_6/ptodata/1/paa/US086.COMB.pep:\*  
11: /cgn2\_6/ptodata/1/paa/US087.COMB.pep:\*  
12: /cgn2\_6/ptodata/1/paa/US088.COMB.pep:\*  
13: /cgn2\_6/ptodata/1/paa/US089.COMB.pep:\*  
14: /cgn2\_6/ptodata/1/paa/US090.COMB.pep:\*  
15: /cgn2\_6/ptodata/1/paa/US091.COMB.pep:\*  
16: /cgn2\_6/ptodata/1/paa/US092.COMB.pep:\*  
17: /cgn2\_6/ptodata/1/paa/US093.COMB.pep:\*  
18: /cgn2\_6/ptodata/1/paa/US094.COMB.pep:\*  
19: /cgn2\_6/ptodata/1/paa/US095.COMB.pep:\*  
20: /cgn2\_6/ptodata/1/paa/US096.COMB.pep:\*  
21: /cgn2\_6/ptodata/1/paa/US097.COMB.pep:\*  
22: /cgn2\_6/ptodata/1/paa/US098.COMB.pep:\*  
23: /cgn2\_6/ptodata/1/paa/US099.COMB.pep:\*  
24: /cgn2\_6/ptodata/1/paa/US100.COMB.pep:\*  
25: /cgn2\_6/ptodata/1/paa/US101.COMB.pep:\*  
26: /cgn2\_6/ptodata/1/paa/US102.COMB.pep:\*  
27: /cgn2\_6/ptodata/1/paa/US60.COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	95.2	9	1	PCT-US99-13975B-130
2	40	95.2	9	21	US-09-701-588B-130
3	40	95.2	9	22	US-09-869-003-4
4	40	95.2	11	22	US-09-869-003-1
5	40	95.2	20	22	US-09-869-003-14
6	40	95.2	29	1	PCT-US99-13975B-136

7 40 95.2 29 1 PCT-US99-13975B-137  
8 40 95.2 29 21 US-09-701-588B-136  
9 40 95.2 29 21 US-09-701-588B-137  
10 40 95.2 29 22 US-09-869-003-31  
11 35 83.0 809 21 US-09-791-537-105715  
12 34 81.0 491 27 US-60-248-498-116  
13 34 81.0 529 27 US-60-248-542-201  
14 34 81.0 1465 27 US-60-360-039-5100  
15 33 78.6 38 1 PCT-US01-00663-30302  
16 33 78.6 38 22 US-09-864-761-37171  
17 33 78.6 38 25 US-10-182-993-29399  
18 33 78.6 38 25 US-10-182-995-23643  
19 33 78.6 38 25 US-10-182-997-22337  
20 33 78.6 38 25 US-10-182-998-13923  
21 33 78.6 38 25 US-10-203-134-30005  
22 33 78.6 38 26 US-10-203-135-28979  
23 33 78.6 38 26 US-10-203-136-30027  
24 33 78.6 38 26 US-10-203-137-30302  
25 33 78.6 38 26 US-10-203-138-14294  
26 33 78.6 38 26 US-10-203-139-29167  
27 33 78.6 38 27 US-60-236-359-19332  
28 33 78.6 192 20 US-09-614-150-10014  
29 33 78.6 192 27 US-60-167-217-10032  
30 33 78.6 192 27 US-60-173-464-8095  
31 33 78.6 192 27 US-60-191-637-10046  
32 33 78.6 192 27 US-60-191-681-7838  
33 33 78.6 272 16 US-09-270-767-45403  
34 33 78.6 559 21 US-09-791-537-137186  
35 33 78.6 571 19 US-09-513-996A-44490  
36 33 78.6 571 20 US-09-570-581A-1807  
37 33 78.6 571 20 US-09-620-394B-4020  
38 33 78.6 218 23 US-09-972-211-80  
39 33 78.6 218 24 US-10-096-625-80  
40 33 78.6 220 23 US-09-972-211-89  
41 33 78.6 220 24 US-10-096-625-89  
42 33 78.6 256 21 US-09-791-537-152655  
43 33 78.6 301 1 PCT-US02-05109-922  
44 33 78.6 301 20 US-09-629-469A-13446  
45 33 78.6 318 26 US-10-219-999-59560

#### ALIGNMENTS

RESULT 1  
PCT-US99-13975B-130  
Sequence 130, Application PC/TUS9913975B  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL, INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
TITLE OF INVENTION: EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
TITLE OF INVENTION: PEPTIDE IMMUNOGENS  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US99/13975B  
FILING DATE: 21-JUNE-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,412  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:

Sequence 137, App  
Sequence 136, App  
Sequence 137, App  
Sequence 31, App  
Sequence 105715, App  
Sequence 116, App  
Sequence 201, App  
Sequence 5100, App  
Sequence 30302, A  
Sequence 37171, A  
Sequence 29399, A  
Sequence 23643, A  
Sequence 22337, A  
Sequence 13923, A  
Sequence 30005, A  
Sequence 28979, A  
Sequence 30027, A  
Sequence 30302, A  
Sequence 14294, A  
Sequence 29167, A  
Sequence 19332, A  
Sequence 10014, A  
Sequence 10032, A  
Sequence 8095, App  
Sequence 10046, A  
Sequence 7838, App  
Sequence 45403, A  
Sequence 137186, A  
Sequence 44490, A  
Sequence 1807, App  
Sequence 4020, App  
Sequence 90, App  
Sequence 90, App  
Sequence 89, App  
Sequence 89, App  
Sequence 152655, App  
Sequence 922, App  
Sequence 13446, A  
Sequence 59560, A

NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US99-139758-130

Query Match 95.2%; Score 40; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

RESULT 2  
US-09-701-588B-130  
Sequence 130, Application US/09701588B  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
PEPTIDE IMMUNOGENS  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/701,588B  
FILING DATE: 29-Nov-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-09-701-588B-130

Query Match 95.2%; Score 40; DB 21; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

Db 1 KSSGKLISL 9

RESULT 3  
US-09-869-003-4  
Sequence 4, Application US/09869003  
GENERAL INFORMATION:  
APPLICANT: Scala, Giuseppe  
APPLICANT: Chen, Xueni  
APPLICANT: Cohen, Oren J.  
APPLICANT: Fauci, Anthony  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
Department of Health and Human Services  
TITLE OF INVENTION: Novel HIV Related Peptides  
FILE REFERENCE: 015280-386200US  
CURRENT APPLICATION NUMBER: US/09/869,003  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: US 60/115,430  
PRIOR FILING DATE: 1999-01-11  
PRIOR APPLICATION NUMBER: US 60/132,760  
PRIOR FILING DATE: 1999-05-06  
PRIOR APPLICATION NUMBER: WO PCT/US00/00372  
PRIOR FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: p195 epitope  
US-09-869-003-4

Query Match 95.2%; Score 40; DB 22; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.2e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

RESULT 4  
US-09-869-003-1  
Sequence 1, Application US/09869003  
GENERAL INFORMATION:  
APPLICANT: Scala, Giuseppe  
APPLICANT: Chen, Xueni  
APPLICANT: Cohen, Oren J.  
APPLICANT: Fauci, Anthony  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
Department of Health and Human Services  
TITLE OF INVENTION: Novel HIV Related Peptides  
FILE REFERENCE: 015280-386200US  
CURRENT APPLICATION NUMBER: US/09/869,003  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: US 60/115,430  
PRIOR FILING DATE: 1999-01-11  
PRIOR APPLICATION NUMBER: US 60/132,760  
PRIOR FILING DATE: 1999-05-06  
PRIOR APPLICATION NUMBER: WO PCT/US00/00372  
PRIOR FILING DATE: 2000-01-07  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: antigenic



FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 10  
OTHER INFORMATION: /note="(e-N)Lys"  
PCT-US99-13975B-137

Query Match 95.2%; Score 40; DB 1; Length 29;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

RESULT 8  
US-09-701-588B-136  
Sequence 136, Application US/09701588B  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
PEPTIDE IMMUNOGENS  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/701,588B  
FILING DATE: 29-NOV-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 136:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 10

OTHER INFORMATION: /note="(e-N)Lys"  
SEQUENCE DESCRIPTION: SEQ ID NO: 136:  
US-09-701-588B-136

Query Match 95.2%; Score 40; DB 21; Length 29;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

RESULT 9  
US-09-701-588B-137  
Sequence 137, Application US/09701588B  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
PEPTIDE IMMUNOGENS  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/701,588B  
FILING DATE: 29-NOV-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PC1  
TELEPHONE: 212-758-4800  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 29 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified site  
LOCATION: 10  
OTHER INFORMATION: /note="(e-N)Lys"  
SEQUENCE DESCRIPTION: SEQ ID NO: 137:  
US-09-701-588B-137

Query Match 95.2%; Score 40; DB 21; Length 29;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
|||||  
Db 1 KSSGKLISL 9

RESULT 10  
US-09-869-003-31  
Sequence 31, Application US/09869003

```
; GENERAL INFORMATION:
; APPLICANT: Scala, Giuseppe
; APPLICANT: Chen, Xueni
; APPLICANT: Cohen, Oren J.
; APPLICANT: Fauci, Anthony
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Novel HIV Related Peptides
; FILE REFERENCE: 015280-386200US
; CURRENT APPLICATION NUMBER: US/09/869, 003
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/115,430
; PRIOR FILING DATE: 1999-01-11
; PRIOR APPLICATION NUMBER: US 60/132,760
; PRIOR FILING DATE: 1999-05-06
; PRIOR APPLICATION NUMBER: WO PCT/US00/00372
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:antigenic
; NAME/KEY: MOD_RES
; LOCATION: (1)..(45)
; OTHER INFORMATION: Xaa = any amino acid not identical to the amino
; OTHER INFORMATION: acid naturally flanking the subsequence at positions
; OTHER INFORMATION: 46-54 in HIV-1, may be present or absent
; NAME/KEY: MOD_RES
; LOCATION: (53)..(99)
; OTHER INFORMATION: Xaa = any amino acid not identical to the amino
; OTHER INFORMATION: acid naturally flanking the subsequence at positions
; OTHER INFORMATION: 46-54 in HIV-1, may be present or absent
US-09-869-003-31

Query Match          95.2%; Score 40; DB 22; Length 99;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
Db 46 KSSGKLISL 54

RESULT 11
US-09-791-537-105715
; Sequence 105715, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 105715
; LENGTH: 809
; TYPE: PRT
; ORGANISM: Hordeum vulgare
US-09-791-537-105715

Query Match          83.3%; Score 35; DB 21; Length 809;
Best Local Similarity 77.8%; Pred. No. 6e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
```

---

```
Db 543 KSDGKLIVSL 551

RESULT 12
US-60-248-498-116
; Sequence 116, Application US/60248498
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PHASE I PROTEINS, AND USES THEREOF
; FILE REFERENCE: C1000910
; CURRENT APPLICATION NUMBER: US/60/248,498
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 116
; LENGTH: 491
; TYPE: PRT
; ORGANISM: HUMAN
US-60-248-498-116

Query Match          81.0%; Score 34; DB 27; Length 491;
Best Local Similarity 77.8%; Pred. No. 5.6e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
Db 423 KNAGKLISL 431

RESULT 13
US-60-248-542-201
; Sequence 201, Application US/60248542
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN PHASE I
; FILE REFERENCE: C1000946
; CURRENT APPLICATION NUMBER: US/60/248,542
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Human
US-60-248-542-201

Query Match          81.0%; Score 34; DB 27; Length 529;
Best Local Similarity 77.8%; Pred. No. 6.1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KSSGKLISL 10
Db 382 KNAGKLISL 390

RESULT 14
US-60-360-039-5100
; Sequence 5100, Application US/60360039
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Chen, Xianfeng
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)A
; CURRENT APPLICATION NUMBER: US/60/360,039
```

```

: CURRENT FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 5100
: LENGTH: 1465
: TYPE: PRT
: ORGANISM: Caenorhabditis elegans
US-60-360-039-5100

```

```

Query Match      81.0%; Score 34; DB 27; Length 1465;
Best Local Similarity 88.9%; Pred. No. 1.9e+03;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 KSSGKLISL 10
      |||||
Db      708 KSSGKLIL 716

```

```

RESULT 15
PCT-US01-00663-30302
: Sequence 30302, Application PC/TUS0100663
: GENERAL INFORMATION:
: APPLICANT: Molecular Dynamics, Inc.
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: PB 0004 WO 7
: CURRENT APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 04 February 2000 (04.02.00)
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 26 May 2000 (26.05.00)
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 03 August 2000 (03.08.00)
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 03 October 2000 (03.10.00)
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 27 September 2000 (27.09.00)
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 21 September 2000 (21.09.00)
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 30 June 2000 (30.06.00)
: NUMBER OF SEQ ID NOS: 38837
: SOFTWARE: Molecular Dynamics Sequence Listing Engine
: SEQ ID NO 30302
: LENGTH: 38
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC009296.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA. SIGNAL = 0.92
: OTHER INFORMATION: EST_HUMAN HIT: BE883574.1, EVALU0 1.10e+00
: OTHER INFORMATION: SWISSPROT HIT: P16415, EVALU0 8.20e-02
PCT-US01-00663-30302

```

```

Query Match      78.6%; Score 33; DB 1; Length 38;
Best Local Similarity 87.5%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2 KSSGKLIS 9
      |||||
Db      17 KSAGKLIS 24

```

Search completed: January 10, 2003, 09:46:27  
 Job time : 381 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2003, 09:37:44 ; Search time 15 seconds  
(without alignments)  
52.294 Million cell updates/sec

Title: US-09-869-003-1  
Perfect score: 42  
Sequence: 1 KSSGKLISLX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 264427 seqs, 71309735 residues

Total number of hits satisfying chosen parameters: 264427

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/1/paa/PCr\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/1/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	95.2	9	US-09-701-588C-130	Sequence 130, App
2	40	95.2	29	US-09-701-588C-136	Sequence 136, App
3	40	95.2	29	US-09-701-588C-137	Sequence 137, App
4	33	78.6	38	US-10-203-138A-14294	Sequence 14294, A
5	31	73.8	443	PCT-US02-33645-8	Sequence 8, Appl
6	31	73.8	443	PCT-US02-33645-12	Sequence 12, Appl
7	31	73.8	445	PCT-US02-33645-4	Sequence 4, Appl
8	31	73.8	1422	US-09-724-676-88052	Sequence 88052, A
9	31	73.8	1422	US-09-724-676-88040	Sequence 88040, A
10	31	73.8	1422	US-09-724-676A-88040	Sequence 88040, A
11	31	73.8	1422	US-09-724-676A-88052	Sequence 88052, A
12	31	73.8	1443	US-09-724-676A-88051	Sequence 88051, A
13	31	73.8	1443	US-09-724-676A-88039	Sequence 88039, A
14	31	73.8	1444	US-09-724-676A-88039	Sequence 88039, A
15	31	73.8	1444	US-09-724-676A-88037	Sequence 88037, A
16	31	73.8	1737	US-09-724-676-88038	Sequence 88038, A
17	31	73.8	1737	US-09-724-676A-88037	Sequence 88037, A
18	31	73.8	1737	US-09-724-676A-88038	Sequence 88038, A
19	31	73.8	4368	US-09-724-676A-73643	Sequence 73643, A
20	31	73.8	4368	US-09-724-676A-73643	Sequence 73643, A
21	31	73.8	238	US-10-203-138A-10552	Sequence 10552, A
22	30	71.4	362	US-09-751-708A-80	Sequence 80, Appl
23	30	71.4	364	US-10-203-138A-10553	Sequence 10553, A
24	30	71.4	69	US-10-203-138A-11172	Sequence 11172, A
25	29	69.0	127	US-10-092-411A-4589	Sequence 4589, Ap

27	29	69.0	334	1	PCT-US02-33541-4	Sequence 4, Appl
28	29	69.0	334	6	US-10-131-813A-12	Sequence 12, Appl
29	29	69.0	334	6	US-10-131-813A-12	Sequence 12, Appl
30	29	69.0	334	6	US-10-131-823A-12	Sequence 12, Appl
31	29	69.0	334	6	US-10-131-823A-12	Sequence 12, Appl
32	29	69.0	334	6	US-10-131-826A-12	Sequence 12, Appl
33	29	69.0	334	6	US-10-131-829A-12	Sequence 12, Appl
34	29	69.0	334	6	US-10-125-926A-12	Sequence 12, Appl
35	29	69.0	334	6	US-10-127-829A-12	Sequence 12, Appl
36	29	69.0	334	6	US-10-127-831A-12	Sequence 12, Appl
37	29	69.0	334	6	US-10-127-835A-12	Sequence 12, Appl
38	29	69.0	334	6	US-10-127-837A-12	Sequence 12, Appl
39	29	69.0	334	6	US-10-127-842A-12	Sequence 12, Appl
40	29	69.0	334	6	US-10-127-850A-12	Sequence 12, Appl
41	29	69.0	334	6	US-10-127-901A-12	Sequence 12, Appl
42	29	69.0	334	6	US-10-128-689A-12	Sequence 12, Appl
43	29	69.0	334	6	US-10-131-830A-12	Sequence 12, Appl
44	29	69.0	334	6	US-10-131-833A-12	Sequence 12, Appl
45	29	69.0	334	6	US-10-131-837A-12	Sequence 12, Appl

## ALIGNMENTS

RESULT 1  
US-09-701-588C-130  
Sequence 130, Application US/09701588C  
GENERAL INFORMATION:  
APPLICANT: UNITED BIOMEDICAL, INC., ET AL.  
TITLE OF INVENTION: ARTIFICIAL T HELPER CELL  
EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC  
NUMBER OF SEQUENCES: 151  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10154-0054  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC Windows  
SOFTWARE: Word 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/701,588C  
FILING DATE: 29-Nov-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/100,414  
FILING DATE: 20-JUNE-1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Maria H. Lin  
REGISTRATION NUMBER: 29,323  
REFERENCE/DOCKET NUMBER: 1151-4158PCI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-6849  
TELEFAX: 212-751-6849  
INFORMATION FOR SEQ ID NO: 130:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-09-701-588C-130  
Query Match 95.2%, Score 40, DB 5, length 9;  
Best Local Similarity 100.0%, Pred. No. 2.5e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0;  
Gaps 0;  
QY 2 KSSGKLISL 10

Db 1 KSSGKLISL 9

## RESULT 2

US-09-701-588C-136

Sequence 136, Application US/09701588C

GENERAL INFORMATION:

APPLICANT: UNITED BIOMEDICAL INC., ET AL.

TITLE OF INVENTION: ARTIFICIAL T HELPER CELL

EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC

PEPTIDE IMMUNOGENS

NUMBER OF SEQUENCES: 151

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/701.588C

FILING DATE: 29-Nov-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100.414

FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4158PC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 136:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified site

LOCATION: 10

OTHER INFORMATION: /note="(e-N)Lys"

SEQUENCE DESCRIPTION: SEQ ID NO: 136:

US-09-701-588C-136

Query Match 95.2%; Score 40; DB 5; Length 29;

Best Local Similarity 100.0%; Pred. No. 0.064;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10

Db 1 KSSGKLISL 9

## RESULT 3

US-09-701-588C-137

Sequence 137, Application US/09701588C

GENERAL INFORMATION:

APPLICANT: UNITED BIOMEDICAL INC., ET AL.

TITLE OF INVENTION: ARTIFICIAL T HELPER CELL

EPITOPES AS IMMUNE STIMULATORS FOR SYNTHETIC

PEPTIDE IMMUNOGENS

NUMBER OF SEQUENCES: 151

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan, L.L.P.

STREET: 345 Park Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10154-0054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC Windows

SOFTWARE: Word 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/701.588C

FILING DATE: 29-Nov-2002

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/100.414

FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:

NAME: Maria H. Lin

REGISTRATION NUMBER: 29,323

REFERENCE/DOCKET NUMBER: 1151-4158PC1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-758-4800

TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO: 137:

SEQUENCE CHARACTERISTICS:

LENGTH: 29 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Modified site

LOCATION: 10

OTHER INFORMATION: /note="(e-N)Lys"

SEQUENCE DESCRIPTION: SEQ ID NO: 137:

US-09-701-588C-137

Query Match 95.2%; Score 40; DB 5; Length 29;

Best Local Similarity 100.0%; Pred. No. 0.064;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10

Db 1 KSSGKLISL 9

## RESULT 4

US-10-203-138A-14294

Sequence 14294, Application US/10203138A

GENERAL INFORMATION:

APPLICANT: Molecular Dynamics, Inc.

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

FILE REFERENCE: PB 0004 WO 8

CURRENT APPLICATION NUMBER: US/10/203.138A

PRIOR FILING DATE: 2002-08-02

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 04 February 2000 (04.02.00)

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 26 May 2000 (26.05.00)

PRIOR APPLICATION NUMBER: US 09/633,366

PRIOR FILING DATE: 03 August 2000 (03.08.00)

PRIOR APPLICATION NUMBER: GB 24263,6

PRIOR FILING DATE: 03 October 2000 (03.10.00)

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 27 September 2000 (27.09.00)

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 21 September 2000 (21.09.00)

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 30 June 2000 (30.06.00)

```

; NUMBER OF SEQ ID NOS: 15438
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 14294
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009296.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.93
; FEATURE:
; OTHER INFORMATION: EST_HUMAN HIT: BE883574.1, EVALUATE 1.10e+00
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: P16415, EVALUATE 8.20e-02
; US-10-203-138A-14294

Query Match          78.6%; Score 33; DB 6; Length 38;
Best Local Similarity 87.5%; Pred. No. 2.6;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 KSSGKLIS 9
DB 17 KSAGKLIS 24

RESULT 5
PCT-US02-33645-8
; Sequence 8, Application PC/TUS0233645
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guangping
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Simian Adenovirus Nucleic Acid and Amino Acid Sequences, Vectors
; FILE REFERENCE: UPN-02677PCT
; CURRENT FILING DATE: 2002-11-20
; CURRENT APPLICATION NUMBER: PCT/US02/33645
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/331,951
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/366,798
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 443
; TYPE: PRT
; ORGANISM: chimpanzee adenovirus serotype Pan6
; PCT-US02-33645-8

Query Match          73.8%; Score 31; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SSGKLIS 9
DB 74 SSGKLIS 80

RESULT 6
PCT-US02-33645-12
; Sequence 12, Application PC/TUS0233645
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guangping
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Simian Adenovirus Nucleic Acid and Amino Acid Sequences, Vectors
; FILE REFERENCE: UPN-02677PCT
; CURRENT FILING DATE: 2002-11-20
; CURRENT APPLICATION NUMBER: PCT/US02/33645
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/331,951
```

```

; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/366,798
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 443
; TYPE: PRT
; ORGANISM: chimpanzee adenovirus serotype Pan7
; PCT-US02-33645-12

Query Match          73.8%; Score 31; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SSGKLIS 9
DB 74 SSGKLIS 80

RESULT 7
PCT-US02-33645-4
; Sequence 4, Application PC/TUS0233645
; GENERAL INFORMATION:
; APPLICANT: The Trustees of the University of Pennsylvania
; APPLICANT: Wilson, James M.
; APPLICANT: Gao, Guangping
; APPLICANT: Roy, Soumitra
; TITLE OF INVENTION: Simian Adenovirus Nucleic Acid and Amino Acid Sequences, Vectors
; FILE REFERENCE: UPN-02677PCT
; CURRENT FILING DATE: 2002-11-20
; CURRENT APPLICATION NUMBER: PCT/US02/33645
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/331,951
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/366,798
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 445
; TYPE: PRT
; ORGANISM: chimpanzee adenovirus serotype Pan5
; PCT-US02-33645-4

Query Match          73.8%; Score 31; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SSGKLIS 9
DB 74 SSGKLIS 80

RESULT 8
US-09-724-676-88040
; Sequence 88040, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 88040
; LENGTH: 1422
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676-88040

Query Match          73.8%; Score 31; DB 5; Length 1422;
Best Local Similarity 77.8%; Pred. No. 4.2e+02;
```

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 9  
US-09-724-676-88052  
; Sequence 88052, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88052  
; LENGTH: 1422  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-88052

Query Match 73.8%; Score 31; DB 5; Length 1422;  
Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 10  
US-09-724-676A-88040  
; Sequence 88040, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88040  
; LENGTH: 1422  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-88040

Query Match 73.8%; Score 31; DB 5; Length 1422;  
Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 11  
US-09-724-676A-88052  
; Sequence 88052, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88052  
; LENGTH: 1422  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-724-676A-88052

Query Match 73.8%; Score 31; DB 5; Length 1422;  
Best Local Similarity 77.8%; Pred. No. 4.2e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 12  
US-09-724-676-88051  
; Sequence 88051, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88051  
; LENGTH: 1443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676-88051

Query Match 73.8%; Score 31; DB 5; Length 1443;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 13  
US-09-724-676A-88051  
; Sequence 88051, Application US/09724676A  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676A  
; CURRENT FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 97222  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 88051  
; LENGTH: 1443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-724-676A-88051

Query Match 73.8%; Score 31; DB 5; Length 1443;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:||||| |  
Db 786 RSSGKLILL 794

RESULT 14  
US-09-724-676-88039  
; Sequence 88039, Application US/09724676  
; GENERAL INFORMATION:  
; APPLICANT: Compugen LTD  
; TITLE OF INVENTION: Variants of alternative splicing  
; FILE REFERENCE: 129181.4 Compugen  
; CURRENT APPLICATION NUMBER: US/09/724,676  
; CURRENT FILING DATE: 2000-11-28

NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO: 88039  
LENGTH: 1444  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676-88039

Query Match 73.8%; Score 31; DB 5; Length 1444;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:|||||  
DB 786 RSSGKLILL 794

RESULT 15  
US-09-724-676A-88039  
Sequence 88039, Application US/09724676A  
GENERAL INFORMATION:  
APPLICANT: Compugen LTD  
TITLE OF INVENTION: Variants of alternative splicing  
FILE REFERENCE: 129181.4 Compugen  
CURRENT APPLICATION NUMBER: US/09/724,676A  
CURRENT FILING DATE: 2000-11-28  
NUMBER OF SEQ ID NOS: 97222  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO: 88039  
LENGTH: 1444  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-724-676A-88039

Query Match 73.8%; Score 31; DB 5; Length 1444;  
Best Local Similarity 77.8%; Pred. No. 4.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KSSGKLISL 10  
:|||||  
DB 786 RSSGKLILL 794

Search completed: January 10, 2003, 09:39:59  
Job time : 16 secs

**THIS PAGE BLANK (USPTO)**

09/869003

FILE: REGISTRY ENTERED AT 14:42:07 ON 10 JAN 2003

L1 5 S KSSGKLISL/SQSP

L1 ANSWER 1 OF 5 REGISTRY COPYRIGHT 2003 ACS

RN 285989-12-0 REGISTRY

CN Peptide, (Xaa-Lys-Ser-Ser-Gly-Lys-Leu-Ile-Ser-Leu-Xaa) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 1: PN: WO0042068 SEQID: 1 claimed protein

CI MAN

SQL 11

SEQ 1 XKSSGKLISL X

=====

HITS AT: 2-10

REFERENCE 1: 133:134165

L1 ANSWER 2 OF 5 REGISTRY COPYRIGHT 2003 ACS

RN 284673-35-4 REGISTRY

CN L-Lysine, L-.alpha.-glutamylglycyl-L-.alpha.-glutamyl-L-phenylalanyl-L-cysteinyl-L-lysyl-L-seryl-L-serylglycyl-L-lysyl-L-leucyl-L-isoleucyl-L-seryl-L-leucyl-L-cysteinylglycyl-L-.alpha.-aspartyl-L-prolyl-L-alanyl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 14: PN: WO0042068 SEQID: 14 claimed protein

SQL 20

SEQ 1 EGEFCKSSGK LISLCGDPK

=====

HITS AT: 6-14

REFERENCE 1: 133:134165

L1 ANSWER 3 OF 5 REGISTRY COPYRIGHT 2003 ACS

RN 253332-95-5 REGISTRY

CN Peptide, (Lys-Ser-Ser-Gly-Lys-Leu-Ile-Ser-Leu-Xaa-Ile-Ser-Ile-Thr-Glu-Ile-Arg-Thr-Val-Ile-Val-Thr-Arg-Ile-Glu-Thr-Ile-Leu-Phe) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 134: PN: WO9966957 SEQID: 137 claimed sequence

CI MAN

SQL 29

SEQ 1 KSSGKLISLX ISITEIRTVI VTRIETILF

=====

HITS AT: 1-9

REFERENCE 1: 132:77607

L1 ANSWER 4 OF 5 REGISTRY COPYRIGHT 2003 ACS

RN 253332-94-4 REGISTRY

CN Peptide, (Lys-Ser-Ser-Gly-Lys-Leu-Ile-Ser-Leu-Xaa-Ile-Ser-Ile-Ser-Glu-Ile-Lys-Gly-Val-Ile-Val-His-Lys-Ile-Glu-Gly-Ile-Leu-Phe) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 39: PN: WO9966957 SEQID: 136 claimed protein

CI MAN

**THIS PAGE BLANK (USPTO)**



09/869003

SQL 29

SEQ 1 KSSGKLISLX ISISEIKGVI VHKIEGILF

=====

HITS AT: 1-9

REFERENCE 1: 132:77607

L1 ANSWER 5 OF 5 REGISTRY COPYRIGHT 2003 ACS

RN 232933-99-2 REGISTRY

CN L-Leucine, L-lysyl-L-seryl-L-serylglycyl-L-lysyl-L-leucyl-L-isoleucyl-L-seryl- (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 127: PN: WO9966957 SEQID: 130 claimed sequence

CN 4: PN: WO0042068 SEQID: 4 claimed protein

SQL 9

SEQ 1 KSSGKLISL

=====

HITS AT: 1-9

REFERENCE 1: 136:100838

REFERENCE 2: 133:134165

REFERENCE 3: 132:77607

REFERENCE 4: 131:115005

FILE 'HCAPLUS' ENTERED AT 14:43:25 ON 10 JAN 2003

L2 4 S L1

L2 ANSWER 1 OF 4 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2001:836703 HCAPLUS

DOCUMENT NUMBER: 136:100838

TITLE: Protection of rhesus macaques against disease progression from pathogenic SHIV-89.6PD by vaccination with phage-displayed HIV-1 epitopes

AUTHOR(S): Chen, Xueni; Scala, Giuseppe; Quinto, Ileana; Liu, Weimin; Chun, Tae-Wook; Justement, J. Shawn; Cohen, Oren J.; vanCott, Tom C.; Iwanicki, Marcin; Lewis, Mark G.; Greenhouse, Jack; Barry, Todd; Venzon, David; Fauci, Anthony S.

CORPORATE SOURCE: Laboratory of Immunoregulation, NIAID, NIH, Bethesda, MD, USA

SOURCE: Nature Medicine (New York, NY, United States) (2001), 7(11), 1225-1231

CODEN: NAMEFI; ISSN: 1078-8956

PUBLISHER: Nature America Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The antigenic polymorphism of HIV-1 is a major obstacle in developing an effective vaccine. Accordingly, we screened random peptide libraries (RPLs) displayed on phage with antibodies from HIV-infected individuals and identified an array of HIV-specific epitopes that behave as antigenic mimics of conformational epitopes of gp120 and gp41 proteins. We report that the selected epitopes

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/869003

are shared by a collection of HIV-1 isolates of clades A-F. The phage-borne epitopes are immunogenic in rhesus macaques, where they elicit envelope-specific antibody responses. Upon i.v. challenge with 60 MID50 of pathogenic SHIV-89.6PD, all monkeys became infected; however, in contrast to the naive and mock-immunized monkeys, four of five mimotope-immunized monkeys experienced lower levels of peak viremia, followed by viral set points of undetectable or transient levels of viremia and a mild decline of CD4+ T cells, and were protected from progression to AIDS-like illness. These results provide a new approach to the design of broadly protective HIV-1 vaccines.

IT 232933-99-2

RL: PAC (Pharmacological activity); THU (Therapeutic use); BIOL (Biological study); USES (Uses)

(protection of rhesus macaques against disease progression from pathogenic SHIV-89.6PD by vaccination with phage-displayed HIV-1 epitopes)

REFERENCE COUNT: 44 THERE ARE 44 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L2 ANSWER 2 OF 4 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2000:493570 HCAPLUS

DOCUMENT NUMBER: 133:134165

TITLE: Novel HIV related peptides

INVENTOR(S): Scala, Giuseppe; Chen, Xueni; Cohen, Oren J.; Fauci, Anthony

PATENT ASSIGNEE(S): United States Dept. of Health and Human Services, USA

SOURCE: PCT Int. Appl., 54 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000042068	A2	20000720	WO 2000-US372	20000107
WO 2000042068	A3	20001221		
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1159297	A2	20011205	EP 2000-904245	20000107
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			

PRIORITY APPLN. INFO.: US 1999-115430P P 19990111  
US 1999-132760P P 19990506  
WO 2000-US372 W 20000107

AB This invention is the discovery of novel specific epitopes and antibodies assocd. with long term survival of HIV-1 infections. These epitopes and antibodies have use in prepg. vaccines for

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/869003

IT preventing HIV-1 infection or for controlling progression to AIDS.  
232933-99-2 284673-35-4 285989-12-0D,  
derivs.  
RL: BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); USES (Uses)  
(novel HIV related peptides for use as vaccine or for inducing  
passive immunity against HIV or HIV-1 infection and progression  
to AIDS)

L2 ANSWER 3 OF 4 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:819268 HCAPLUS

DOCUMENT NUMBER: 132:77607

TITLE: Artificial T helper cell epitopes as immune  
stimulators for synthetic peptide immunogens

INVENTOR(S): Wang, Chang Yi

PATENT ASSIGNEE(S): United Biomedical Inc., USA

SOURCE: PCT Int. Appl., 129 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 9966957	A2	19991229	WO 1999-US13975	19990621
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
AU 9945808	A1	20000110	AU 1999-45808	19990621
BR 9912177	A	20010410	BR 1999-12177	19990621
EP 1089760	A1	20010411	EP 1999-928826	19990621
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			
JP 2002518463	T2	20020625	JP 2000-555643	19990621
PRIORITY APPLN. INFO.:			US 1998-100412	A2 19980620
			WO 1999-US13975	W 19990621

AB Disclosed are immunogenic peptide compns. comprising a T helper epitope linked to a target antigenic site or a synthetic B cell epitope, and optionally with a immunostimulatory sequence for inducing T helper cell-mediated immune response and producing high level of antibodies directed against the "target antigen". The disclosed immunogenic peptide compns. are useful for contraception or control of hormone-dependent tumor (with LH-releasing hormone as target), growth promotion in farm animal (with somatostatin as target), treatment of allergy (with IgE as target), prevention of HIV infection (with CD4 receptor as target), prevention of foot-and-mouth disease (with FMDV capsid protein as target), treatment of malaria (with circumsporozoite antigen of Plasmodium falciparum as target), and treatment of arteriosclerosis (with cholesteryl ester transport protein as target).

IT 232933-99-2 253332-94-4 253332-95-5

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/869003

RL: BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); USES (Uses)  
(artificial T helper cell epitopes as immune stimulators for  
synthetic peptide immunogens)

L2 ANSWER 4 OF 4 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 1999:329703 HCAPLUS

DOCUMENT NUMBER: 131:115005

TITLE: Selection of HIV-specific immunogenic epitopes  
by screening random peptide libraries with  
HIV-1-positive sera

AUTHOR(S): Scala, Giuseppe; Chen, Xueni; Liu, Weimin;  
Telles, Jean Noel; Cohen, Oren J.; Vaccarezza,  
Mauro; Igarashi, Tatsu; Fauci, Anthony S.

CORPORATE SOURCE: Laboratory of Immunoregulation, National  
Institute of Allergy and Infectious Diseases,  
National Institutes of Health, Bethesda, MD,  
20892, USA

SOURCE: Journal of Immunology (1999), 162(10), 6155-6161  
CODEN: JOIMA3; ISSN: 0022-1767

PUBLISHER: American Association of Immunologists

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Efforts to develop a protective HIV-1 vaccine have been hindered by  
difficulties in identifying epitopes capable of inducing broad  
neutralizing Ab responses. In fact, the high mutation rate  
occurring in HIV-1 envelope proteins and the complex structure of  
gp120 as an oligomer assocd. with gp41 result in a high degree of  
antigenic polymorphism. To overcome these obstacles, the authors  
screened random peptide libraries using sera from HIV-infected  
subjects to identify antigenic and immunogenic mimics of HIV-1  
epitopes. After extensive counter-screening with HIV-neg. sera, the  
authors isolated peptides specifically recognized by Abs from  
HIV-1-infected individuals. These peptides behaved as antigenic  
mimics of linear or conformational HIV-1 epitopes generated in vivo  
in infected subjects. Consistent with these findings, sera of  
simian HIV-infected monkeys also recognized the HIV-specific  
epitopes. The selected peptides were immunogenic in mice, where  
they elicited HIV-specific Abs that effectively neutralized HIV-1  
isolates. These results demonstrate the pools of HIV-1 mimotopes  
can be selected from combinatorial peptide libraries by taking  
advantage of the HIV-specific Ab repertoire induced by the natural  
infection.

IT 232933-99-2

RL: BAC (Biological activity or effector, except adverse); BSU  
(Biological study, unclassified); PRP (Properties); BIOL (Biological  
study)

(selection of immunodeficiency virus-specific immunogenic  
epitopes by screening with HIV-1-pos. sera)

REFERENCE COUNT: 38 THERE ARE 38 CITED REFERENCES AVAILABLE  
FOR THIS RECORD. ALL CITATIONS AVAILABLE  
IN THE RE FORMAT

FILE 'HOME' ENTERED AT 14:43:38 ON 10 JAN 2003

**THIS PAGE BLANK (USPTO)**